



Уральский  
федеральный  
университет

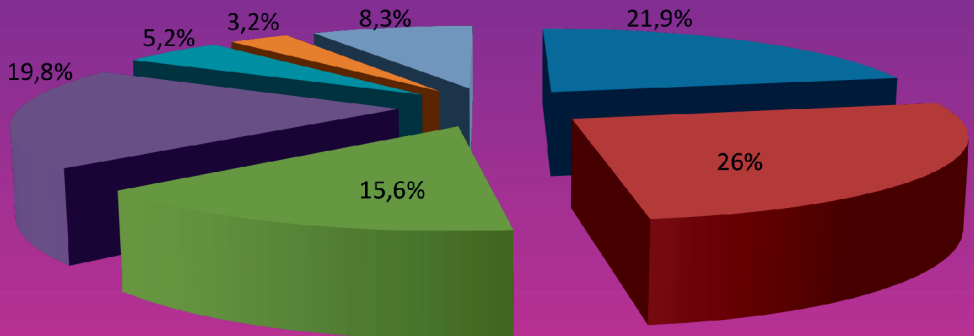
имени первого Президента  
России Б.Н.Ельцина

Уральский  
энергетический  
институт

**М. А. ПЛЕСКУНОВ**

# МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА СОЦИОЛОГИЧЕСКИХ ДАННЫХ

Учебное пособие





Министерство образования и науки Российской Федерации

Уральский федеральный университет  
имени первого Президента России Б. Н. Ельцина

**М. А. Плескунов**

## **МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА СОЦИОЛОГИЧЕСКИХ ДАННЫХ**

Рекомендовано методическим советом  
Уральского федерального университета  
в качестве **учебного пособия** для студентов вуза,  
обучающихся по направлению  
39.04.03 — Организация работы с молодежью

Екатеринбург  
Издательство Уральского университета  
2017

УДК 303.7(075.8)

ББК 60.6я73

ПЗ8

Рецензенты:

кафедра «Высшая и прикладная математика» УрГУПС (зав. кафедрой проф., д-р физ.-мат. наук *Г. А. Тимофеева*); вед. науч. сотр. ИММ УрО РАН, д-р физ.-мат. наук *Ю. И. Бердышев*

Научный редактор — проф., д-р физ.-мат. наук *А. Н. Сесекин*  
**Плескунов, М. А.**

ПЗ8 Методы статистического анализа социологических данных : учеб. пособие / М. А. Плескунов. — Екатеринбург : Изд-во Урал. ун-та, 2017. — 144 с.  
ISBN 978-5-7996-1988-6

Содержит теоретический материал и примеры применения методов статистического анализа к задачам социологических исследований. Даны приемы обработки и графического отражения первичных данных опроса исследуемой социальной группы, методы исследования наличия зависимости между рядами данных и теснотой их связи и статистические методы проверки социологических гипотез.

Библиогр.: 20 назв.

УДК 303.7(075.8)

ББК 60.6я73

ISBN 978-5-7996-1988-6

© Уральский федеральный  
университет, 2017

# ОСНОВНЫЕ ПОНЯТИЯ СТАТИСТИКИ

---

**Статистика** — наука о методах сбора, обработки и количественного анализа эмпирических данных.

**Генеральная совокупность** — совокупность объектов, составляющих предмет изучения.

**Выборочная совокупность** — часть объектов генеральной совокупности, выбранная для непосредственного исследования и подвергшаяся испытанию (опросу, измерению, наблюдению и т. п.).

**Статистическая сводка** — совокупность эмпирических данных статистических испытаний, полученных часто из различных источников и собранных воедино.

**Выборочный метод** — метод, позволяющий делать убедительные (с определенной вероятностью) выводы о свойствах всей генеральной совокупности на основе изучения выборочной совокупности.

**Доверительный интервал** — интервал, содержащий (покрывающий) с заданной надежностью интересующий нас неизвестный параметр.

**Надежность** (доверительная вероятность) — вероятность того, что интересующая нас неизвестная величина находится в определенном интервале.

**Закон распределения** — правило (формула), позволяющее установить функцию распределения вероятностей случайной величины.

**Уровень значимости** — вероятность совершить ошибку первого рода.

**Ошибка первого рода** — ошибка, заключающаяся в том, что в результате статистической проверки будет отвергнута гипотеза, являющаяся на самом деле верной.

**Ошибка второго рода** — ошибка, заключающаяся в том, что в результате статистической проверки будет принята основная гипотеза, в то время как на самом деле верной является альтернативная гипотеза.

**Номинальная переменная** — переменная, измеренная в номинальной шкале, т.е. в шкале наименований. Устанавливается лишь принадлежность значений переменной определенным классам.

**Ранговая переменная** — переменная, измеренная в ранговой, т.е. в порядковой, шкале. Значения переменной упорядочены по предпочтению.

**Количественная переменная** — переменная, измеренная в метрической шкале. Каждому значению переменной поставлено в соответствие определенное число.

**Качественная переменная** — переменная, измеренная в номинальной или ранговой шкале. Каждое значение переменной имеет свое наименование (в номинальной шкале) или имеет свой ранг (в ранговой шкале).

**Альтернативная переменная** — допускает респонденту выбор только одного варианта ответа.

**Поливариантная переменная** — позволяет респонденту выбрать несколько вариантов ответа одновременно.

**Дискретная переменная** — переменная, множество значений которой представляет конечную или бесконечную последовательность отдельных, изолированных друг от друга значений.

**Непрерывная переменная** — переменная, значения которой заполняют сплошь некоторый промежуток числовой оси или же всю числовую ось.

**Ограниченная переменная** — переменная, имеющая лишь конечное число возможных значений (все их можно перечислить).

**Неограниченная переменная** — переменная, число значений которой ничем не ограничено, т. е. их бесконечно много (перечислить все их невозможно).

**Дисперсия** — мера рассеяния случайной величины относительно ее среднего значения.

**Среднеквадратическое отклонение** — чаще всего используемая мера отклонения случайной величины от ее среднего значения. Среднеквадратическое отклонение равно корню квадратному из дисперсии. В отличие от дисперсии имеет ту же размерность, что и сама случайная величина.

**Альтернативный (дихотомический) признак** — признак, для которого фиксируется только его наличие или отсутствие у объекта исследования.

**Вариативный признак** — признак, имеющий определенную градацию своих значений, т. е. обладающий упорядоченной последовательностью значений.

**Вариационный ряд** — упорядоченная по возрастанию или по убыванию последовательность значений признака.

**Мода** — наиболее часто встречающееся значение переменной.

**Медиана** — середина вариационного ряда.

**Функциональная зависимость** — однозначная зависимость между величинами, когда каждому значению одной (независимой) величины соответствует одно определённое значение другой (зависимой, результирующей) величины.

**Корреляционная зависимость** или **корреляция** — вероятностная или статистическая зависимость, не имеющая строго функционального характера. В отличие от функциональной, корреляционная зависимость возникает тогда, когда один из признаков зависит не только от данного второго, но и от ряда случайных факторов или же когда среди условий, от которых зависят и тот и другой признаки, имеются общие для них обоих условия. В силу этого каждому значению одной величины соответствует некоторое множество возможных значений другой; какое именно из них она примет, зависит от иных факторов (случайных или отличающих эти величины друг от друга).

**Регрессия** — зависимость среднего значения какой-либо величины от некоторой другой величины или от нескольких величин (в теории вероятностей и математической статистике).

**Уравнение регрессии**  $Y$  по  $X$  — уравнение  $y = u(x)$ , которое отражает изменение величины  $Y$  в среднем при изменении  $x$  (значений независимой переменной  $X$ ); соответствующий такому уравнению график называется линией регрессии величины  $Y$  по  $X$ .

**Фактор** — причина, движущая сила или условие какого-либо процесса, определяющая его характер или отдельные его черты.

**Факторный анализ** — метод, применяемый для установления и изучения взаимосвязей между значениями переменных; позволяет описать объект измерения всесторонне и в то же время компактно.

**Кластер** — объединение однородных элементов, которое может рассматриваться как самостоятельная единица, обладающая определёнными свойствами; класс родственных элементов статистической совокупности.

**Кластерный анализ** — многомерная статистическая процедура, выполняющая сбор данных, содержащих информацию о выборке объектов, и затем упорядочивающая объекты в сравнительно однородные группы (кластеры).



Выбор статистических показателей зависит от вида используемых данных и метода их измерений. В шкалах более высокого уровня можно использовать все показатели, которые используются в шкалах более низкого уровня, но не все показатели, используемые в шкалах более высокого уровня можно использовать в шкалах более низкого уровня.

Применяемые статистические показатели в зависимости от типа использованной шкалы измерений

Тип шкалы	Статистические показатели
Номинальная	Количество опрошенных, количество ответивших, количество ответов. Частоты, проценты, доли. Мода. Дисперсия доли
Порядковая	Количество опрошенных, количество ответивших, количество ответов. Частоты, проценты, доли, накопленные частоты. Мода, медиана. Дисперсия доли
Метрическая	Количество опрошенных, количество ответивших, количество ответов. Частоты, проценты, доли, накопленные частоты. Мода, медиана, среднее арифметическое. Размах вариации, стандартное отклонение, коэффициент вариации. Асимметрия, эксцесс. Дисперсия, дисперсия доли

## ФОРМУЛЫ СТАТИСТИКИ

---

### Обозначения основных показателей

Показатель	Генеральная совокупность	Выборочная совокупность
Объем	$N$	$n$
Значение признака	$X_i$	$x_i$
Частота	$N_i$	$n_i$ или $f_i$
Относительная частота	$W_i = \frac{N_i}{N}$	$w_i = \frac{n_i}{n} = \frac{f_i}{n}$
Среднее значение	$\bar{X}$	$\bar{x}$
Доля единиц, обладающих данным значением альтернативного признака	$p$	$w$
Доля единиц, не обладающих данным значением альтернативного признака	$q = 1 - p$	$1 - w$
Средняя доля	$\bar{p} = p$	$\bar{w} = w$
Дисперсия вариативного (количественного) признака	$D = \sigma_{\text{ген}}^2$	$\sigma^2$
Дисперсия альтернативного признака (дисперсия доли)	$pq$	$w(1 - w)$

Исправленная выборочная дисперсия вариативного (количественного) признака	–	$s^2 = \frac{n}{n-1} \sigma^2$
Исправленная выборочная дисперсия альтернативного признака (доли)	–	$\frac{n}{n-1} w(1-w)$
Дисперсия среднего значения	$\frac{D}{n} = \frac{\sigma_{\text{ген}}^2}{n}$	$\frac{\sigma^2}{n}$
Дисперсия средней доли	$\frac{pq}{n}$	$\frac{w(1-w)}{n}$
Среднеквадратическое отклонение вариативного (количественного) признака	$\sigma_{\text{ген}} = \sqrt{D}$	$\sigma$
Среднеквадратическое отклонение альтернативного признака (доли)	$\sqrt{pq}$	$\sqrt{w(1-w)}$
Среднеквадратическое отклонение среднего значения	$\sqrt{\frac{D}{n}} = \sqrt{\frac{\sigma_{\text{ген}}^2}{n}}$	$\sqrt{\frac{\sigma^2}{n}}$
Среднеквадратическое отклонение средней доли	$\sqrt{\frac{pq}{n}}$	$\sqrt{\frac{w(1-w)}{n}}$

### Средние величины

Величины	Для не сгруппированных данных	Для сгруппированных данных
Среднее арифметическое	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	$\bar{x} = \frac{1}{n} \sum_{i=1}^m x_i n_i$
Среднее гармоническое	$\bar{x}_{\text{гар}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$	$\bar{x}_{\text{гар}} = \frac{n}{\sum_{i=1}^m \frac{n_i}{x_i}}$

Величины	Для не сгруппированных данных	Для сгруппированных данных
Среднее геометрическое	$\bar{x}_{\text{геом}} = \sqrt[n]{x_1 x_2 \dots x_n}$	$\bar{x}_{\text{геом}} = \sqrt[n_1 + n_2 + \dots + n_m]{x_1^{n_1} x_2^{n_2} \dots x_m^{n_m}}$
Среднее квадратическое	$\bar{x}_{\text{кв}} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$	$\bar{x}_{\text{кв}} = \sqrt{\frac{1}{n} \sum_{i=1}^m x_i^2 n_i}$
Среднее значение квадратов показателя	$\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$	$\overline{x^2} = \frac{1}{n} \sum_{i=1}^m x_i^2 n_i$

### Структурные средние

Показатель	Для дискретного вариационного ряда	Для интервального вариационного ряда (с интервалами одинаковой длины)
Мода	$M_o = x_k$ , где $x_k$ такое, что $n_k = \max_i n_i$	$M_o = x_0 + h \frac{n_{M_o} - n_{M_o-1}}{(n_{M_o} - n_{M_o-1}) + (n_{M_o} - n_{M_o+1})}$
Медиана	$M_e = x_{\frac{n+1}{2}}$ , если $n$ нечетное; $M_e = \frac{1}{2} \left( x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right)$ , если $n$ четное	$M_e = x_0 + h \frac{\frac{n}{2} - S_{Me-1}}{n_{Me}}$

*Примечание:*  $x_0$  — нижняя граница модального (медианного) интервала,  $h$  — длина модального (медианного) интервала,  $n$  — объем выборки,  $n_{M_o}$  — частота модального интервала,  $n_{M_o-1}$  — частота предмодального интервала,  $n_{M_o+1}$  — частота постмодального интервала,  $n_{Me}$  — частота медианного интервала,  $S_{Me-1} = \sum_{i=1}^{Me-1} n_i$  — накопленная частота предмедианного интервала.

Формула Стерджесса для определения оптимального числа групп данных:  $m = 1 + 3,322 \lg n$ .

Совокупность считается однородной, если коэффициент вариации не превышает 33 % (для распределений, близких к нормальным).

Между средним линейным и средним квадратическим отклонениями существует примерное соотношение  $\sigma \approx 1,25\bar{d}$ , если фактическое распределение близко к нормальному.

### Меры вариации (колеблемости)

Показатель	Для несгруппированных данных	Для сгруппированных данных
Размах вариации	$R = x_{\max} - x_{\min}$	
Коэффициент осцилляции	$V_R = \frac{R}{\bar{x}} 100\%$	
Дисперсия	$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$	$\sigma^2 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^2 n_i$
	$\sigma^2 = \overline{x^2} - (\bar{x})^2$ , где	
	$\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$	$\overline{x^2} = \frac{1}{n} \sum_{i=1}^m x_i^2 n_i$
Среднеквадратическое отклонение	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^2 n_i}$
Коэффициент вариации	$V_\sigma = \frac{\sigma}{\bar{x}} 100\%$	
Среднее линейное отклонение	$\bar{d} = \frac{1}{n} \sum_{i=1}^n  x_i - \bar{x} $	$\bar{d} = \frac{1}{n} \sum_{i=1}^m  x_i - \bar{x}  n_i$
Коэффициент линейной вариации	$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} 100\%$	

### *Поправка Шеппарда для выборочной дисперсии интервального вариационного ряда*

Для вычисления дисперсии интервального вариационного ряда каждый интервал заменяется его серединой. В этом случае при расчете выборочной дисперсии возникает систематическая ошибка. Чтобы уменьшить эту ошибку, вводят поправку Шеппарда и находят уточненное значение выборочной дисперсии:  $\sigma_{\text{ут}}^2 = \sigma^2 - \Delta^2 / 12$ , где  $\Delta$  — длина интервала (для равновеликих интервалов).

### *Доверительные интервалы*

Доверительные интервалы для среднего значения:  $(\bar{x} - \Delta_{\bar{x}}, \bar{x} + \Delta_{\bar{x}})$ , доверительные интервалы для средней доли:  $(w - \Delta_p, w + \Delta_p)$ .

Большая выборка ( $n \geq 30$ )

Показатель	Повторный отбор	Бесповторный отбор
Средняя ошибка вариативного признака	$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}$	$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
Средняя ошибка альтернативного признака	$\mu_p = \sqrt{\frac{w(1-w)}{n}}$	$\mu_p = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$
Предельная ошибка вариативного признака	$\Delta_{\bar{x}} = t \sqrt{\frac{\sigma^2}{n}}$	$\Delta_{\bar{x}} = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
Предельная ошибка альтернативного признака	$\Delta_p = t \sqrt{\frac{w(1-w)}{n}}$	$\Delta_p = t \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

Показатель	Повторный отбор	Бесповторный отбор
Необходимый объем выборки для вариативного признака	$n \geq \frac{t^2 \sigma^2}{\Delta_x^2}$	$n \geq \frac{t^2 \sigma^2 N}{t^2 \sigma^2 + N \Delta_x^2}$
Необходимый объем выборки для альтернативного признака	$n \geq \frac{t^2 w(1-w)}{\Delta_p^2}$	$n \geq \frac{t^2 w(1-w)N}{t^2 w(1-w) + N \Delta_p^2}$

Коэффициенты доверия для доверительных интервалов нормального распределения (корни уравнения  $\Phi(t) = \frac{\gamma}{2}$ )

$\gamma$	0,683	0,799	0,899	0,950	0,954	0,990	0,997	0,999
$t$	1,00	1,28	1,64	1,96	2,00	2,58	3,00	3,29

Примечание:  $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{z^2}{2}} dz$  — функция Лапласа.

Малая выборка ( $n < 30$ )

Показатель	Повторный отбор	Бесповторный отбор
Средняя ошибка вариативного признака	$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n-1}}$	$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n-1} \left(1 - \frac{n}{N}\right)}$
Средняя ошибка альтернативного признака*	$\mu_p = \sqrt{\frac{w(1-w)}{n-1}}$	$\mu_p = \sqrt{\frac{w(1-w)}{n-1} \left(1 - \frac{n}{N}\right)}$
Предельная ошибка вариативного признака	$\Delta_{\bar{x}} = t_{\gamma} \sqrt{\frac{\sigma^2}{n-1}}$	$\Delta_{\bar{x}} = t_{\gamma} \sqrt{\frac{\sigma^2}{n-1} \left(1 - \frac{n}{N}\right)}$
Предельная ошибка альтернативного признака*	$\Delta_p = t_{\gamma} \sqrt{\frac{w(1-w)}{n-1}}$	$\Delta_p = t_{\gamma} \sqrt{\frac{w(1-w)}{n-1} \left(1 - \frac{n}{N}\right)}$

Показатель	Повторный отбор	Бесповторный отбор
Необходимый объем выборки для вариативного признака	$n \geq \frac{t_{\gamma}^2 \sigma^2}{\Delta_x^2} + 1$	$n \geq \frac{(t_{\gamma}^2 \sigma^2 + \Delta_x^2) N}{t_{\gamma}^2 \sigma^2 + N \Delta_x^2}$
Необходимый объем выборки для альтернативного признака*	$n \geq \frac{t_{\gamma}^2 w(1-w)}{\Delta_p^2} + 1$	$n \geq \frac{(t_{\gamma}^2 w(1-w) + \Delta_p^2) N}{t_{\gamma}^2 w(1-w) + N \Delta_p^2}$

Примечание: \*формулы применяют при условии:  $nw > 5$  и  $n(1-w) > 5$ .

Значения  $t_{\gamma} = t(\gamma, n)$

$n \backslash \gamma$	0,95	0,99	0,999	$n \backslash \gamma$	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	$\infty$	1,960	2,576	3,291
19	2,10	2,88	3,92				



## Корреляционно-регрессионный анализ

### Линейная регрессия

Уравнение линейной регрессии	$y = ax + b$
Угловой коэффициент уравнения линейной регрессии	$a = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x^2}$
Свободный член уравнения линейной регрессии	$b = \bar{y} - a\bar{x}$
Коэффициент корреляции Пирсона	$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y} = a \frac{\sigma_x}{\sigma_y}$
Коэффициент детерминации	$r^2 = \frac{\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \left( \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y} \right)^2 = a^2 \frac{\sigma_x^2}{\sigma_y^2},$ <p>где <math>\tilde{y}_i = ax_i + b</math></p>
Отклонение	$\varepsilon_i = y_i - \tilde{y}_i$

Коэффициент детерминации показывает величину вариации переменной  $y$ , которая объясняется переменной  $x$ , при наличии линейной связи этих величин. В случае строгой функциональной линейной зависимости между переменными  $x$  и  $y$  коэффициент детерминации  $r^2 = 1$ . Если линейная зависимость между  $x$  и  $y$  отсутствует, то  $r^2 = 0$ .

$\sum_{i=1}^n (y_i - \bar{y})^2$  — это общая вариация переменной  $y$ .

$\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2$  — это вариация переменной  $y$ , которая объясняется

формулой  $y = ax + b$ :

$$y_i - \bar{y} = y_i - \tilde{y}_i + \tilde{y}_i - \bar{y} = (\tilde{y}_i - \bar{y}) + (y_i - \tilde{y}_i) = (\tilde{y}_i - \bar{y}) + \varepsilon_i .$$

$$\sum_{i=1}^n (y_i - \tilde{y}_i)^2 = \sum_{i=1}^n \varepsilon_i^2 \quad \text{— это вариация переменной } y, \text{ которая}$$

не объясняется формулой  $y = ax + b$ . Разница  $y_i - \tilde{y}_i = \varepsilon_i$  называется ошибкой (остатком, отклонением). Значения коэффициентов  $a$  и  $b$  в уравнении  $y = ax + b$  подбираются (методом наименьших квадратов) так, чтобы минимизировать сумму квадратов всех отклонений:  $\sum_{i=1}^n \varepsilon_i^2 \rightarrow \min$ .

### Определение силы связи по значению коэффициента корреляции

#### Теснота связи

Значение коэффициента корреляции	Сила связи
$r = 0$	Линейная связь отсутствует
$0 <  r  \leq 0,3$	Слабая линейная связь
$0,3 <  r  \leq 0,5$	Линейная связь средняя
$0,5 <  r  \leq 0,8$	Линейная связь сильная
$0,8 <  r  < 1$	Очень сильная линейная связь
$ r  = 1$	Имеется функциональная линейная зависимость $y$ от $x$

Чем ближе модуль  $r$  к 1, тем ближе связь между  $x$  и  $y$  к линейной. При  $r = 0$  линейной связи между  $x$  и  $y$  не существует (но, возможно, между  $x$  и  $y$  существует другая (не линейная) зависимость).

Направление связи оценивается по знаку коэффициента корреляции. Положительный коэффициент корреляции говорит о наличии прямой связи, а отрицательный — о наличии обратной связи. Знак коэффициента корреляции совпадает со знаком углового коэффициента уравнения линейной регрессии:  $ra \geq 0$ .

### *Проверка значимости коэффициента корреляции Пирсона*

Для вывода о наличии линейной связи между переменными  $x$  и  $y$  в генеральной совокупности проверяют гипотезу о значимости найденного коэффициента Пирсона в выборочной совокупности.

Основная гипотеза —  $H_0 : r_r = 0$ . Альтернативная гипотеза —  $H_1 : r_r \neq 0$ .

Двусторонняя критическая область: граничные точки  $\pm t_{\alpha; n-2}$ , где  $t_{\alpha; n-2}$  — критическая точка распределения Стьюдента для двусторонней области при уровне значимости  $\alpha$  и числе степеней свободы  $k = n - 2$ .

Статистика  $t = \sqrt{r^2(n-2) / (1-r^2)}$ .

Если  $t > t_{\alpha; n-2}$ , то нулевая гипотеза  $H_0$  отклоняется, и принимается альтернативная гипотеза  $H_1$  на уровне значимости  $\alpha$ . В генеральной совокупности есть линейная связь между переменными  $x$  и  $y$ .

Если  $t < t_{\alpha; n-2}$ , то нет оснований отвергать нулевую гипотезу об отсутствии линейной связи между переменными в генеральной совокупности.

### *Коэффициент ранговой корреляции Спирмена*

Применяется для оценки степени связи двух качественных признаков, измеренных в ранговой шкале. Направление связи оценивается по знаку коэффициента ранговой корреляции. Поло-

жительный коэффициент корреляции говорит о наличии прямой связи, а отрицательный — о наличии обратной связи.

$$r_s = 1 - 6 \frac{\sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

где  $d_i$  — разница между парой рангов,  $n$  — количество сравниваемых пар рангов.

$-1 \leq r_s \leq 1$ . При  $r_s = 1$  имеется полная прямая связь (зависимость) признаков, при  $r_s = -1$  имеется полная обратная связь (зависимость) признаков. Чем ближе к нулю  $|r_s|$ , тем зависимость признаков меньше.

### *Проверка значимости коэффициента ранговой корреляции Спирмена (при количестве сравниваемых пар $n \geq 9$ )*

Основная гипотеза —  $H_0 : r_s = 0$ . Альтернативная гипотеза —  $H_1 : r_s \neq 0$ .

$T_{кр} = t_{\alpha; n-2} \sqrt{(1 - r_s^2) / (n - 2)}$ , где  $t_{\alpha; n-2}$  — критическая точка распределения Стьюдента (для двусторонней критической области). Если  $|r_s| < T_{кр}$ , то нет оснований отвергать нулевую гипотезу, т. е. связь между признаками незначима. Если  $|r_s| > T_{кр}$ , то нулевая гипотеза отвергается на уровне значимости  $\alpha$  (с вероятностью  $\alpha$  ошибки первого рода). Между признаками существует значимая ранговая корреляционная связь.

### *Коэффициент ранговой корреляции Кендалла*

Коэффициент ранговой корреляции Кендалла (при отсутствии связанных рангов, т. е. элементов с одинаковыми рангами) имеет

те же свойства, что и коэффициент Спирмена, и применяется аналогично.

$$r_k = \frac{4R}{n(n-1)} - 1,$$

где  $n$  — количество сравниваемых пар,  $R = \sum_{i=1}^{n-1} R_i$ ,  $R_i$  — число ран-

гов, больших ранга  $i$ -го элемента второго ряда, расположенных правее этого элемента, при условии, что ряды упорядочены по возрастанию номера ранга (в порядке ухудшения качества признака) элементов первого ряда. Для проверки значимости коэффициента Кендалла используют критическую точку

$T_{кр} = t\sqrt{2(2n+5)/(9n(n-1))}$ , где  $\Phi(t) = \gamma/2$ ,  $\gamma = 1 - \alpha$  для уровня значимости  $\alpha$  (см. стр. 3). Если  $|r_k| < T_{кр}$ , то нет оснований отвергать нулевую гипотезу, т.е. связь между признаками незначима. Если  $|r_k| > T_{кр}$ , то между признаками существует значимая ранговая корреляционная связь.

## Асимметрия и эксцесс

### Центральные моменты

Показатель	Для несгруппированных данных	Для сгруппированных данных
Центральный момент первого порядка	$\mu_1 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) = 0$	$\mu_1 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x}) n_i = 0$
Центральный момент второго порядка	$\mu_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \sigma^2$	$\mu_2 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^2 n_i = \sigma^2$
Центральный момент третьего порядка	$\mu_3 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3$	$\mu_3 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^3 n_i$

Показатель	Для несгруппированных данных	Для сгруппированных данных
Центральный момент четвертого порядка	$\mu_4 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4$	$\mu_4 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^4 n_i$
Центральный момент $k$ -го порядка	$\mu_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$	$\mu_k = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^k n_i$

Асимметрия	$A_s = \frac{\mu_3}{\sigma^3}$	При $A_s < 0$ — левосторонняя асимметрия, при $A_s > 0$ — правосторонняя асимметрия
Экспесс	$E_k = \frac{\mu_4}{\sigma^4} - 3$	При $E_k < 0$ — плосковершинное распределение, при $E_k > 0$ — островершинное распределение

При левосторонней асимметрии  $M_o > M_e > \bar{x}$ , при правосторонней асимметрии  $M_o < M_e < \bar{x}$ . Мода, медиана и среднее значение связаны примерным соотношением  $\bar{x} - M_e \approx \frac{1}{3}(\bar{x} - M_o)$ .

При  $|A_s| > 0,5$  асимметрия считается значительной, при  $|A_s| < 0,25$  — незначительной.

Для нормального распределения  $\bar{x} = M_o = M_e$ ,  $A_s = 0$ ,  $E_k = 0$ .

#### Среднеквадратические ошибки

Среднеквадратическая ошибка асимметрии	$\sigma_{A_s} = \sqrt{\frac{6(n-1)}{(n+1)(n+3)}}$
Среднеквадратическая ошибка эксцесса	$\sigma_{E_k} = \sqrt{\frac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}}$

В случае  $|A_s|/\sigma_{A_s} > 3$  асимметрия существенна, и распределение признака в генеральной совокупности несимметрично.

### Формулы Линдберга для упрощенной оценки асимметрии и эксцесса

Показатели асимметрии и эксцесса позволяют сделать вывод о возможности отнесения данного эмпирического распределения к типу кривых нормального распределения и, следовательно, правомерности использования при обработке статистических данных законов нормального распределения.

Эксцесс рассчитывается, обычно, для симметричных распределений. Наличие выраженной асимметрии или значительного отрицательного эксцесса может указывать на качественную неоднородность исследуемой совокупности.

$$A_s \approx P_{x > \bar{x}} - 50,$$

где  $P_{x > \bar{x}}$  — доля (в процентах) количества тех вариантов, которые превосходят среднее значение, в общем количестве вариантов данного ряда; 50 — доля (в процентах) количества вариантов, превосходящих среднее значение, в общем количестве вариантов ряда нормального распределения.

$$E_k \approx P_{|x - \bar{x}| < \sigma/2} - 38,29,$$

где  $P_{|x - \bar{x}| < \sigma/2}$  — доля (в процентах) количества вариантов, лежащих в интервале, равном половине среднеквадратического отклонения (в ту или другую сторону от среднего значения) в общем количестве вариантов данного ряда; 38,29 — доля (в процентах) количества вариантов, лежащих в интервале, равном половине среднеквадратического отклонения (в ту или другую сторону от среднего значения) в общем количестве вариантов ряда нормального распределения.

Другие используемые упрощенные оценки асимметрии:

$$A_3 \approx \frac{\bar{x} - M_o}{\sigma} \quad \text{или} \quad A_3 \approx \frac{\bar{x} - M_e}{\sigma}.$$

### Коэффициент Фехнера

Коэффициент Фехнера характеризует наличие статистической связи (зависимости) между двумя переменными ( $x$  и  $y$ ) и ее направление:

$$K_{\Phi} = \frac{a - b}{a + b},$$

где  $a$  — число совпадений знаков у  $\Delta x_i$  и  $\Delta y_i$  ( $i=1, 2, \dots, n$ ),  $b$  — число несовпадений знаков у этих величин.  $\Delta x_i = x_i - \bar{x}$ ,  $\Delta y_i = y_i - \bar{y}$ ,  $\bar{x} = \sum x_i / n$ ,  $\bar{y} = \sum y_i / n$ .

Коэффициент Фехнера может принимать значения от  $-1$  до  $1$ . Это наиболее простой показатель тесноты зависимости между двумя переменными.

### Статистическая проверка гипотез

Гипотеза о среднем значении нормально распределенного признака

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критические точки
$H_0: \bar{X} = a$	$H_1: \bar{X} < a$	левосторонняя	$k_{кр} = -t_{\alpha; n-1}$
$H_0: \bar{X} = a$	$H_1: \bar{X} > a$	правосторонняя	$k_{кр} = t_{\alpha; n-1}$
$H_0: \bar{X} = a$	$H_1: \bar{X} \neq a$	двусторонняя	$k_{кр} = \pm t_{\alpha; n-1}$
Статистика	$t = \frac{\bar{x} - a}{\sigma / \sqrt{n-1}}$		

*Примечание:* Используется распределение Стьюдента.



Гипотеза о генеральной доле нормально распределенного признака

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критические точки
$H_0 : p = a$	$H_1 : p < a$	левосторонняя	$k_{кр} = -z_{\alpha}$
$H_0 : p = a$	$H_1 : p > a$	правосторонняя	$k_{кр} = z_{\alpha}$
$H_0 : p = a$	$H_1 : p \neq a$	двусторонняя	$k_{кр} = \pm z_{\alpha/2}$
Статистика	$z = \frac{w - a}{\sqrt{a(1-a)/n}}$		

*Примечание:* Используется нормальное распределение.

Гипотеза о равенстве генеральных дисперсий  
нормальных совокупностей

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критическая точка
$H_0 : \sigma_{1\text{ген}}^2 = \sigma_{2\text{ген}}^2$	$H_0 : \sigma_{1\text{ген}}^2 \neq \sigma_{2\text{ген}}^2$	двусторонняя	$k_{кр} = F_{\alpha/2; n_A-1; n_B-1}$
Оценки	$s_1^2 = \frac{n}{n-1} \sigma_1^2$ — оценка $\sigma_{1\text{ген}}^2$ ; $s_2^2 = \frac{n}{n-1} \sigma_2^2$ — оценка $\sigma_{2\text{ген}}^2$		
Обозначения	$s_A$ — большая оценка дисперсий, $s_B$ — меньшая оценка дисперсий, $n_A$ — объем выборки, имеющей большую оценку генеральной дисперсии, $n_B$ — объем выборки, имеющей меньшую оценку генеральной дисперсии		
Статистика	$F = s_A / s_B$		

*Примечание:* Используется распределение Фишера.

Гипотеза о равенстве средних значений нормальных совокупностей в случае равенства генеральных дисперсий

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критические точки
$H_0 : a_1 = a_2$	$H_1 : a_1 < a_2$	левосторонняя	$k_{кр} = -t_{\alpha; n_1 + n_2 - 2}$
$H_0 : a_1 = a_2$	$H_1 : a_1 > a_2$	правосторонняя	$k_{кр} = t_{\alpha; n_1 + n_2 - 2}$
$H_0 : a_1 = a_2$	$H_1 : a_1 \neq a_2$	двусторонняя	$k_{кр} = \pm t_{\alpha/2; n_1 + n_2 - 2}$
Статистика	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2 - 2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$		

*Примечание:* Используется распределение Стьюдента.

Гипотеза о равенстве средних значений нормальных совокупностей в случае неравенства генеральных дисперсий  
(при  $n_1 \geq 30$  и  $n_2 \geq 30$ )

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критические точки
$H_0 : a_1 = a_2$	$H_1 : a_1 < a_2$	левосторонняя	$k_{кр} = -z_{\alpha}$
$H_0 : a_1 = a_2$	$H_1 : a_1 > a_2$	правосторонняя	$k_{кр} = z_{\alpha}$
$H_0 : a_1 = a_2$	$H_1 : a_1 \neq a_2$	двусторонняя	$k_{кр} = \pm z_{\alpha/2}$
Статистика	$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1 - 1} + \frac{\sigma_2^2}{n_2 - 1}}}$		

*Примечание:* Используется нормальное распределение.

Гипотеза о равенстве генеральных долей нормальных совокупностей для выборок большого объема ( $n_1 \geq 30$  и  $n_2 \geq 30$ )

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критические точки
$H_0 : p_1 = p_2$	$H_1 : p_1 < p_2$	левосторонняя	$k_{кр} = -z_\alpha$
$H_0 : p_1 = p_2$	$H_1 : p_1 > p_2$	правосторонняя	$k_{кр} = z_\alpha$
$H_0 : p_1 = p_2$	$H_1 : p_1 \neq p_2$	двусторонняя	$k_{кр} = \pm z_{\alpha/2}$
Статистика	$z = \frac{w_1 - w_2}{\sqrt{\tilde{w}(1 - \tilde{w})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ <p>где <math>\tilde{w}</math> — выборочная доля для объединенной выборки</p>		

*Примечание:* Используется нормальное распределение.

Гипотеза о средней разности между парными измерениями нормально распределенных признаков

Основная гипотеза	Альтернативная гипотеза	Критическая область	Критические точки
$H_0 : \bar{X}_d = 0$	$H_1 : \bar{X}_d < 0$	левосторонняя	$k_{кр} = -t_{\alpha; n-1}$
$H_0 : \bar{X}_d = 0$	$H_1 : \bar{X}_d > 0$	правосторонняя	$k_{кр} = t_{\alpha; n-1}$
$H_0 : \bar{X}_d = 0$	$H_1 : \bar{X}_d \neq 0$	двусторонняя	$k_{кр} = \pm t_{\alpha; n-1}$
Статистика	$t = \frac{\bar{x}_d}{\sigma_d / \sqrt{n-1}},$ <p>где <math>d_i = x_{1i} - x_{2i}</math>, <math>\bar{x}_d = \frac{\sum d_i}{n}</math>, <math>\sigma_d = \sqrt{\frac{\sum d_i^2}{n} - (\bar{x}_d)^2}</math></p>		

*Примечание:* Используется распределение Стьюдента.

Гипотеза о принадлежности нового наблюдения нормальной генеральной совокупности

Основная гипотеза	$H_0 : x_{\text{нов}}$ принадлежит той же генеральной совокупности, что и выборка
Альтернативная гипотеза	$H_1 : x_{\text{нов}}$ не принадлежит той же генеральной совокупности, что и выборка
Критическая область	двусторонняя
Критические точки	$k_{\text{кр}} = \pm t_{\alpha; n-1}$
Статистика	$t = \frac{x_{\text{нов}} - \bar{x}}{\sigma \sqrt{(n+1)/(n-1)}},$ <p>где <math>x_{\text{нов}}</math> — результат нового наблюдения</p>

*Примечание:* Используется распределение Стьюдента.

Гипотеза об отсутствии связи между порядковыми данными ( $n \geq 10$ )

Основная гипотеза	$H_0$ : между двумя последовательностями нет связи, они не согласованы друг с другом
Альтернативная гипотеза	$H_1$ : между двумя последовательностями существует какая-то связь
Критическая область	правосторонняя
Критическая точка	$k_{\text{кр}} = z_{\alpha}$
Статистика	$z =  r_s  \sqrt{n-1}, \text{ где } r_s = 1 - 6 \frac{\sum d_i^2}{n(n^2-1)} - \text{ран-}$ <p>говый коэффициент корреляции Спирмена, <math>d_i = x_i - y_i</math></p>

*Примечание:* Используется нормальное распределение.

Критические точки нормального распределения  
(корни уравнения  $P(z > z_\alpha) = \alpha$ ,  $\alpha = 1 - \gamma$ ;  
или соответствующего ему уравнения  $\Phi(z_\alpha) = \gamma - 0,5$ )

$\alpha$	0,001	0,005	0,010	0,015	0,020	0,025	0,030	0,035
$z_\alpha$	3,090	2,576	2,326	2,170	2,054	1,960	1,881	1,812
$\alpha$	0,040	0,045	0,050	0,10	0,15	0,20	0,25	0,40
$z_\alpha$	1,751	1,695	1,645	1,282	1,036	0,842	0,675	0,253

### Критерии согласия

Значения  $\lambda_\alpha$

$\alpha$	0,20	0,10	0,05	0,02	0,01	0,001
$\lambda_\alpha$	1,073	1,224	1,358	1,520	1,627	1,950

Примечание:  $\lambda$  — критерий Колмогорова — Смирнова.

Гипотеза о распределении непрерывной  
случайной величины ( $n \geq 50$ )

Основная гипотеза	$H_0$ : случайная величина имеет распределение близкое к распределению с функцией $F(x)$
Альтернативная гипотеза	$H_1$ : эмпирическое распределение существенно отличается от распределения $F(x)$
Критическая область	правосторонняя
Критическая точка	$k_{кр} = \lambda_\alpha$
Статистика	$\lambda = \max_{x_j}  F(x_j) - F_3(x_j)  \sqrt{n}$

Гипотеза о принадлежности двух выборок одной  
и той же генеральной совокупности (  $n_1 \geq 50$  ,  $n_2 \geq 50$  )

Основная гипотеза	$H_0$ : обе выборки извлечены из одной и той же генеральной совокупности
Альтернативная гипотеза	$H_1$ : выборки извлечены из разных совокупностей
Критическая область	правосторонняя
Критическая точка	$k_{кр} = \lambda_{\alpha}$
Статистика	$\lambda = \max_{x_i}  F_{1\alpha}(x_i) - F_{2\alpha}(x_i)  \sqrt{n} ,$ <p>где <math>n = n_1 n_2 / (n_1 + n_2)</math></p>

# ТИПОВЫЕ ЗАДАЧИ СОЦИОЛОГИЧЕСКОГО АНАЛИЗА

---

## *Задача № 1*

### **Анализ одномерных распределений**

**Цель:** отработать навыки математической обработки и анализа массива социологических данных.

**Задание.** Рассмотрите одномерное распределение, представленное в таблице. Рассчитайте статистические показатели, пригодные для данных, замеренных на этом уровне.

Порядок выполнения:

- 1) определить вид переменной и шкалу измерений;
- 2) определить показатели, пригодные для данной шкалы;
- 3) осуществить расчет выбранных показателей;
- 4) провести социологический анализ показателей.

### ***Пример выполнения работы***

Личностные черты, оцениваемые как привлекательные\*<sup>1</sup>

Личностные черты	Абсолютные значения
Волевые черты	93
Усидчивость, трудолюбие	58
Доброта, чуткость	36
Коммуникативные навыки	33

<sup>1</sup> Исходные данные для задач, помеченных знаком \*, взяты из [20].

Личностные черты	Абсолютные значения
Интеллектуальные черты	29
Эмоциональная устойчивость	16
Внешние данные	14
Организованность	12
Нравственные черты	12
Независимость	11
Профессионализм	9
Ответственность, преданность	6
Обаяние, способность оказывать влияние на других	5
Хорошие манеры	5
Оптимизм	4
Чувство юмора	4
Творческий потенциал (оригинальность, инициативность и др.)	4
Чувство собственного достоинства	3
Открытость	2
Практичность (жадность и др.)	2
Жизненный опыт	1
Итого	359

Количество опрошенных — 466 человек.

Не ответили на данный вопрос 195 человек.

1. Распределение одномерное. Переменная ограниченная, поливариантная, номинальная. Шкала измерений — номинальная. Данные уже отсортированы по частоте упоминаний, что делает таблицу более наглядной.

2. Статистические показатели, соответствующие номинальной шкале: количество опрошенных, количество ответивших, количество ответов, частоты, проценты, мода, дисперсия доли.

3. Расчет показателей.

Количество опрошенных: 466.



Количество ответивших:  $466 - 195 = 271$ .

Количество ответов: 359.

Процент ответивших:  $\frac{271}{466} \cdot 100 = 58,15\%$ .

Подсчет процентов для признака «Волевые черты»:

— от опрошенных:  $\frac{93}{466} = 0,199570815 \approx 0,200 = 20,0\%$ ;

— от ответивших:  $\frac{93}{271} = 0,343173431 \approx 0,343 = 34,3\%$ ;

— от ответов:  $\frac{93}{359} = 0,259052924 \approx 0,259 = 25,9\%$ .

Для остальных признаков — аналогично.

Для анализа выбираем процент от опрошенных, т. к. вопрос обращен ко всем респондентам.

Дисперсию доли подсчитаем по формуле  $\sigma^2 = w(1 - w)$ , где  $w$  — доля выборов признака от числа опрошенных.

Для признака «Волевые черты»:

$\sigma^2 = 0,200(1 - 0,200) = 0,2 \cdot 0,8 = 0,16$ . Здесь  $w = 20,0\% = 0,200$ .

Для остальных признаков подсчет ведется аналогично.

Расчет предельной ошибки выборки проведем по формуле

$$\Delta = t \sqrt{\frac{\sigma^2}{n}},$$

где  $n$  — число опрошенных,  $t$  — коэффициент доверия, зависящий от выбираемой надежности (доверительной вероятности)  $\gamma$ . Его выбирают по таблице значений функций Лапласа  $\Phi(t)$  в соответствии с равенством  $\gamma = 2\Phi(t)$ .

Значения функции Лапласа  $\Phi(t)$

$\gamma$	0,683	0,950	0,954	0,990	0,997
$t$	1,00	1,96	2,00	2,58	3,00

Для надежности 0,950 (95 %) коэффициент доверия  $t = 1,96$ .

Возьмем именно это значение.

Тогда для признака «Волевые черты» предельная ошибка выборки

$$\Delta = 1,96 \sqrt{\frac{0,16}{466}} = 0,036318098 \approx 0,036 = 3,6\%.$$

Предельные ошибки выборки для остальных признаков найдем точно так же. Все расчетные данные внесем в таблицу.

Личностные черты	Частоты	Проценты			Дисперсия доли	Ошибка выборки
		От опро- шенных	От отве- тивших	От ответов		
Волевые черты	93	20,0 %	34,3 %	25,9 %	0,160	3,6 %
Усидчивость, трудолю- бие	58	12,4 %	21,4 %	16,2 %	0,109	3,0 %
Доброта, чуткость	36	7,7 %	13,3 %	10,0 %	0,071	2,4 %
Коммуникативные на- выки	33	7,1 %	12,2 %	9,2 %	0,066	2,3 %
Интеллектуальные черты	29	6,2 %	10,7 %	8,1 %	0,058	2,2 %
Эмоциональная устойчи- вость	16	3,4 %	5,9 %	4,5 %	0,033	1,6 %
Внешние данные	14	3,0 %	5,2 %	3,9 %	0,029	1,5 %
Организованность	12	2,6 %	4,4 %	3,3 %	0,025	1,4 %
Нравственные черты	12	2,6 %	4,4 %	3,3 %	0,025	1,4 %
Независимость	11	2,4 %	4,1 %	3,0 %	0,023	1,4 %
Профессионализм	9	1,9 %	3,3 %	2,5 %	0,019	1,2 %
Ответственность, пре- данность	6	1,3 %	2,2 %	1,7 %	0,013	1,0 %
Обаяние, способность оказывать влияние на других	5	1,1 %	1,8 %	1,4 %	0,011	0,9 %

Личностные черты	Частоты	Проценты			Дисперсия доли	Ошибка выборки
		От опрошенных	От ответивших	От ответов		
Хорошие манеры	5	1,1 %	1,8 %	1,4 %	0,011	0,9 %
Оптимизм	4	0,9 %	1,5 %	1,1 %	0,009	0,8 %
Чувство юмора	4	0,9 %	1,5 %	1,1 %	0,009	0,8 %
Творческий потенциал (оригинальность, инициативность и др.)	4	0,9 %	1,5 %	1,1 %	0,009	0,8 %
Чувство собственного достоинства	3	0,6 %	1,1 %	0,8 %	0,006	0,7 %
Открытость	2	0,4 %	0,7 %	0,6 %	0,004	0,6 %
Практичность (жадность и др.)	2	0,4 %	0,7 %	0,6 %	0,004	0,6 %
Жизненный опыт	1	0,2 %	0,4 %	0,3 %	0,002	0,4 %
Итого	359	77,1 %	132,4 %	100 %	—	—

#### 4. Социологический анализ полученных результатов.

«Волевые черты» стоят на первом месте по частоте упоминания в качестве привлекательных личностных черт. Этот признак назвали 20 % всех опрошенных, 34,3 % всех ответивших на вопрос анкеты; он составляет 25,9 % от всех полученных ответов. Признак, имеющий наибольшую частоту, называется модой. Следовательно, в данном случае мода — «волевые черты». Ошибку выборки можно использовать для анализа значимости различий в количестве выборов между альтернативами. Так, разность между количеством выборов альтернатив «волевые черты» и «усидчивость, трудолюбие» составляет  $20 \% - 12,4 \% = 7,6 \%$ , а суммарная ошибка выборки составляет  $3,6 \% + 3 \% = 6,6 \%$ . Значит, действительно, различие имеет место с надежностью 95 % (в нашем случае),

и в генеральной совокупности первую характеристику выбирают чаще, чем вторую. Если различие в количестве выборов меньше суммарной ошибки репрезентативности (суммарной ошибки выборки), то это может означать, что такая разность возникла в результате применения выборочного метода исследования, и нельзя с уверенностью утверждать, что такое же предпочтение существует и в генеральной совокупности. Так, например, обстоит дело с характеристиками «усидчивость, трудолюбие» и «доброта, чуткость». Разница в количестве выборов для них составляет  $12,4\% - 7,7\% = 4,7\%$ , а суммарная ошибка выборки равна  $3,0\% + 2,4\% = 5,4\%$ . Значит, гарантировать, что и в генеральной совокупности будет существовать такое же предпочтение, какое выявилось среди опрошенных, невозможно.

Аналогичный сравнительный анализ можно провести и для других пар признаков.

## Задача № 2

### Измерение, группировка, отображение и анализ первичных социометрических данных

**Цель:** отработать навыки математической обработки и анализа массива социологических данных.

**Задание.** Определить тип переменной и упорядочить данную совокупность показателей социологического исследования, построить частотный вариационный ряд, найти среднее значение переменной, моду и медиану, построить полигон и эмпирическую функцию распределения.

#### Пример выполнения работы

Ниже приведены данные исследования группы студентов по структуре использования ими свободного времени. На вопрос о количестве посещений дискотек в последний месяц были получены следующие ответы:

2 0 4 2 3 2 1 3 4 3 2 1 0 3 4 3 2 3 1 3 0 1 2 3 4  
3 2 3 0 2 3 2 4 1 0 1 2 3 2 0 2 2 3 0 0 3 3 2 3 0  
0 2 3 0 2 4 1 3 1 4 0 2 1 2 0 4 3 4 0 2 2 0 1 2 3  
4 1 3 4 3 4 1 4 2 0 0 4 1 1 1 0 0 0 1 4 2 0 2 1 2

1. Переменная количественная, дискретного типа. Ее значения: 0, 1, 2, 3, 4.

2. Выделим пять групп, соответствующих значениям переменной. Проведем группировку данных по частотам:

№ группы	1	2	3	4	5
Признак группы, $x_i$	0	1	2	3	4
Частота признака, $f_i$	21	17	25	22	15

Дальнейшие расчеты приведем в следующей табличной форме:

№ групп	Признак группы, $x_i$	Частота признака, $f_i$	$x_i f_i$	Кумулятивная частота	Относительная кумулятивная частота
1	0	21	0	21	0,21
2	1	17	17	38	0,38
3	2	25	50	63	0,63
4	3	22	66	85	0,85
5	4	15	60	100	1,00
$\Sigma$	—	100	193	—	—

3. Найдем среднюю арифметическую по формуле:

$$\bar{x} = \frac{\sum_{i=1}^5 x_i f_i}{\sum_{i=1}^5 f_i}.$$

$$\bar{x} = \frac{193}{100} = 1,93$$

4. Модой называется значение признака с наибольшей частотой. В нашем случае — это значение 2. Значит,  $M_o = 2$ , т.е. чаще всего студенты посещали дискотеку два раза в месяц.

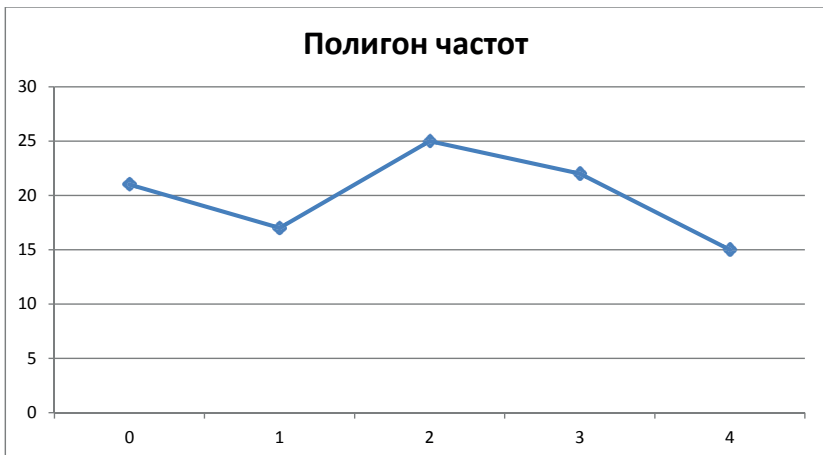
5. Медианой называется значение признака, которое делит вариационный ряд на две равные по количеству элементов части. Найдем номер медианного элемента по формуле  $N_{M_o} = \frac{n+1}{2} = \frac{100+1}{2} = 50,5$ , где  $n$  — объем выборки.

Так как объем выборочной совокупности четный (100 респондентов), то этот номер оказался дробным. Это значит, что медиану нужно рассчитать как среднее арифметическое 50-го и 51-го элементов ранжированного ряда:

$$M_e = \frac{x_{50} + x_{51}}{2} = \frac{2+2}{2} = 2.$$

Первая кумулятивная (накопленная) частота, большая или равная 50, является частотой третьей группы (63). Значит, 50-й и 51-й элемент находятся в этой группе: оба они равны двум. Следовательно, медианой является значение второго признака, т. е. половина опрошенных студентов посещала дискотеку не более двух раз в месяц, а другая половина — не менее двух раз в месяц.

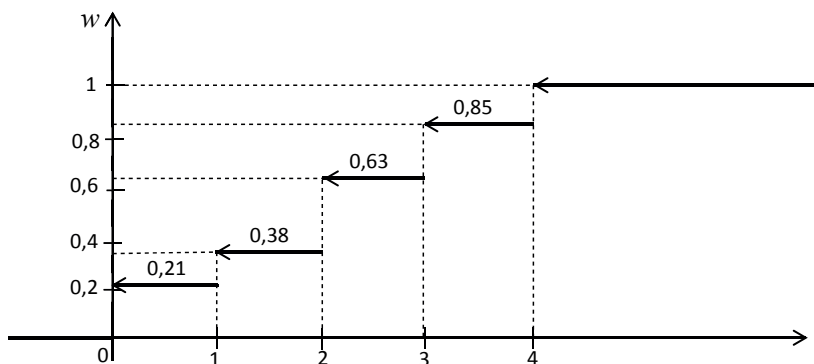
6. Строим полигон частот.



7. Находим эмпирическую функцию распределения

$$w = F_{\mathfrak{g}}(x) = \sum_{x_i < x} w_i = \begin{cases} 0 & \text{при } x \leq 0, \\ 0,21 & \text{при } 0 < x \leq 1, \\ 0,38 & \text{при } 1 < x \leq 2, \\ 0,63 & \text{при } 2 < x \leq 3, \\ 0,85 & \text{при } 3 < x \leq 4, \\ 1 & \text{при } x > 4, \end{cases}$$

и строим ее график:



Здесь  $w$  — относительная частота признака  $w_i = \frac{f_i}{n}$ , где  $f_i$  — частота  $i$ -го значения признака,  $n = \sum f_i$  — объем выборки,

$$w = F_{\text{э}}(x) = \sum_{x_i < x} w_i.$$

Значение эмпирической функции распределения  $w = F_{\text{э}}(x) = \sum_{x_i < x} w_i$  — это накопленная (кумулятивная) относительная частота тех значений признака, которые меньше взятого значения  $x$ . Например, для  $x = 2,5$   $F_{\text{э}}(x) = 0,63$ . Это значит, что суммарная относительная частота (доля) тех, кто посещал дискотеку менее 2,5 раз в месяц, равна 0,63, т.е. 63 %.



### Задача № 3

#### **Измерение, группировка, отображение и анализ первичных социометрических данных в интервальной форме**

**Цель:** отработать навыки группировки, математической обработки и графического отображения первичных социометрических данных в интервальной форме.

**Задание.** Сгруппировать данные социологического исследования, разбив их совокупность на несколько интервалов, построить интервальный вариационный ряд, найти размах вариации, среднее значение переменной, моду и медиану, построить гистограмму и кумуляту, полигон и эмпирическую функцию распределения.

#### ***Пример выполнения работы***

Ниже приведены данные исследования группы студентов, живущих с родителями, по структуре использования ими свободного времени. Студентам было поручено в течение недели хронометрировать и усреднить время, уделяемое ими в день ведению домашнего хозяйства. Были получены следующие ответы (в мин):

27	14	20	16	28	8	5	14	9	23	4	19	16	8	6	14	5	27	16
13	9	17	9	16	14	16	17	23	18	6	22	12	3	18	24	12	8	18
8	15	6	12	6	23	25	4	18	21	17	12	15	9	9	20	21	13	21
5	7	21	10	18	11	13	8	26	6	11	26	32	11	23	19	16	23	14
12	23	11	15	13	22	27	15	4	27	14	13	14	10	28	9	13	20	17

Проанализируем данные по указанной схеме.

1. Найдем минимальный и максимальный элемент совокупности:

$$x_{\min} = 3 \text{ мин}, x_{\max} = 32 \text{ мин.}$$

2. Найдем размах вариации:  $R = x_{\max} - x_{\min} = 32 - 3 = 29 \text{ мин.}$

3. Подавляющее большинство данных находится в промежутке от 5 до 25 мин, поэтому разобьем весь массив данных на шесть ин-

тервалов: до 5, от 5 до 10, от 10 до 15, от 15 до 20, от 20 до 25, 25 мин и больше. Построим вариационный ряд:

Время, мин	До 5	5–10	10–15	15–20	20–25	25 и больше
Число студентов	4	20	24	21	16	10

Верхнюю границу интервала считаем исключенной, а нижнюю — включенной, т.е. если число попадает на границу между двумя интервалами, то его относим не к предыдущему, а к последующему интервалу.

Например, число 10 отнесем к интервалу 10–15, т.е. будем понимать интервал  $a-b$  как  $[a, b)$ .

4. Дальнейшие расчеты оформим в виде таблицы.

№ группы	Интервал $a_i-b_i$	Середина интервала, $x'_i$	Частота, $f_i$	$x'_i f_i$	Накопленная частота
1	0–5	2,5	4	10	4
2	5–10	7,5	20	150	24
3	10–15	12,5	24	300	48
4	15–20	17,5	21	367,5	69
5	20–25	22,5	16	360	85
6	25–30	27,5	10	275	95
Сумма		—	95	1462,5	—

Для удобства последующих расчетов полагаем длину первого интервала (с неопределенной нижней границей) равной длине следующего за ним интервала, т.е., в нашем случае, равной пяти. Длину последнего интервала (с неопределенной верхней границей) считаем равной длине предшествующего интервала, т.е. тоже равной пяти. Таким образом, получаем все интервалы равной длины.

5. Найдем среднее значение признака по формуле средней арифметической взвешенной:  $\bar{x} = \frac{\sum x'_i f_i}{\sum f_i} = \frac{1462,5}{95} = 15,39$  мин.

Таким образом, в среднем студенты этой группы тратят на ведение домашнего хозяйства около 15,5 мин в день.

6. Интервал, имеющий наибольшую частоту, называется модальным интервалом. Моду, т.е. признак с наибольшей частотой, для интервальных вариационных рядов находят по формуле:

$$M_o = x_0 + \Delta \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

где  $x_0$  — нижняя граница модального интервала;  $\Delta$  — длина интервала;  $f_{M_o}$  — частота модального интервала;  $f_{M_o-1}$  — частота предмодального интервала;  $f_{M_o+1}$  — частота постмодального интервала.

В нашем случае модальным интервалом является третий интервал, его нижняя граница  $x_0 = 10$ , длина интервала  $\Delta = 5$ , частота модального интервала  $f_{M_o} = 24$ , частота предмодального интервала  $f_{M_o-1} = 20$ , а постмодального —  $f_{M_o+1} = 21$ . Подставим эти значения в формулу:

$$M_o = 10 + 5 \frac{24 - 20}{(24 - 20) + (24 - 21)} = 10 + 5 \frac{4}{4 + 3} \approx 12,86.$$

Итак, модой является значение 12,86 мин.

Таким образом, чаще всего студенты этой группы тратят на ведение домашнего хозяйства в среднем около 13 мин в день.

7. Медианой называется значение признака, которое делит вариационный ряд на две равные по числу элементов части. Медианным называется интервал, содержащий медиану. Чтобы найти медианный интервал, вычислим номер медианного элемента по формуле  $N_{Me}^0 = \frac{n+1}{2}$ , где  $n$  — объем выборки. Если число  $n$  — четное, то  $N_{Me}^0$  будет дробным числом. Медианным является пер-

вый от начала отсчета интервал, накопленная частота которого больше или равна  $N_{Me}$ , т.е. номеру медианного элемента. Медиана для интервального вариационного ряда вычисляется по формуле:

$$M_e = x_0 + \Delta \frac{n/2 - S_{M_e-1}}{f_{M_e}},$$

где  $x_0$  — нижняя граница медианного интервала;  $\Delta$  — длина интервала;  $f_{M_e}$  — частота медианного интервала;  $S_{M_e-1}$  — накопленная частота предмедианного интервала;  $n = \sum f_i$  — объем выборки.

В нашем случае номер медианного элемента  $N_{Me} = \frac{n+1}{2} = \frac{95+1}{2} = 48$ .

Первый интервал, накопленная частота которого больше или равна номеру медианного элемента, — третий. Это и есть медианный интервал. Его нижняя граница  $x_0 = 10$ , длина интервала  $\Delta = 5$ , частота медианного интервала  $f_{M_e} = 24$ , накопленная частота предмедианного интервала  $S_{M_e-1} = 24$ , объем выборки  $n = 95$ . Подставим эти значения в формулу:

$$M_e = 10 + 5 \frac{95/2 - 24}{24} = 10 + 5 \frac{47,5 - 24}{24} \approx 14,896 \approx 14,90.$$

Итак, медианой данного распределения является значение 14,90.

Таким образом, половина опрошенных студентов тратит на ведение домашнего хозяйства в среднем не более примерно 15 мин в день, а другая половина — не менее 15 мин.

Подведем итоги исследования. Выяснилось, что данная группа студентов в среднем занимается ведением домашнего хозяйства примерно 15,5 мин в день, причем чаще всего встречаются те, кто тратит на это около 13 мин в день, при этом половина студентов в среднем тратит не более 15 мин в день.

8. Отобразим графически полученные данные. Найдем относительные частоты интервалов:  $w_i = \frac{f_i}{n}$ , где  $n = \sum f_i$  — объем выбор-

ки. Найдем также плотности частот интервалов:  $\frac{f_i}{\Delta}$ , где  $\Delta$  — длина интервала.

В нашем случае  $\Delta = 5$ .

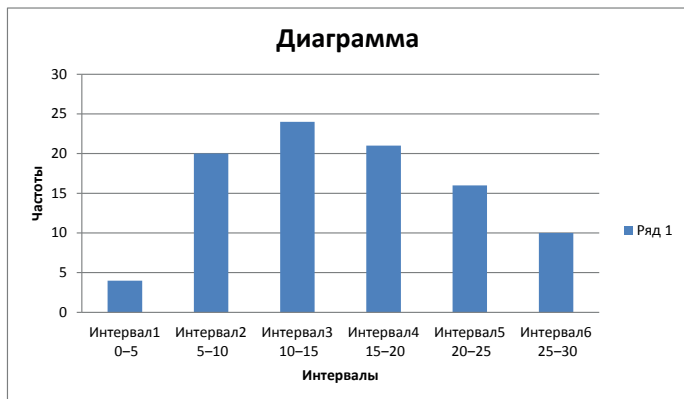
Расчетные данные отразим в табличной форме.

Интервал $a_i - b_i$	Середина интервала, $x'_i$	Частота, $f_i$	Плотность частоты	Относительная частота	Накопленная относительная частота
0–5	2,5	4	0,8	0,042	0,042
5–10	7,5	20	4,0	0,211	0,253
10–15	12,5	24	4,8	0,253	0,506
15–20	17,5	21	4,2	0,221	0,727
20–25	22,5	16	3,2	0,168	0,895
25–30	27,5	10	2,0	0,105	1,000
Сумма	—	95	—	1	—

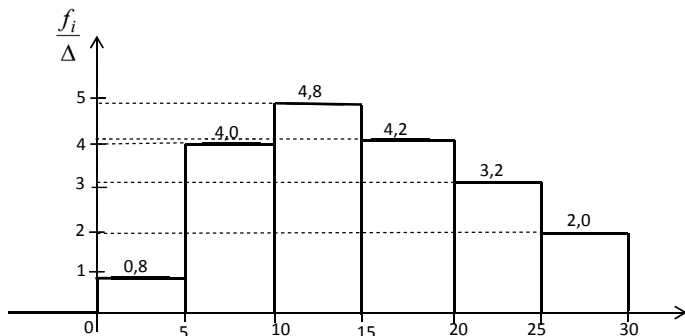
9. Для построения полигона частот соединим соседние точки с координатами  $(x'_i, f_i)$  отрезками прямых.



Эти же данные можно отразить и с помощью столбиковой диаграммы:



10. Для построения гистограммы отложим на горизонтальной оси интервалы, а на вертикальной оси — плотности частот.



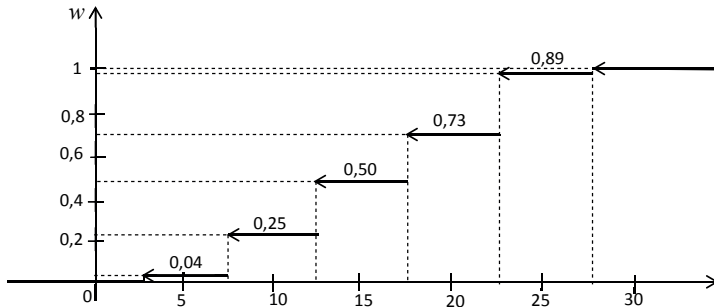
Площадь полученной фигуры равна объему выборки.

11. Для построения эмпирической функции распределения  $F_3$  используем накопленные относительные частоты. Эмпирическая функция распределения строится для середин интервалов, которые рассматриваются как значения дискретной переменной (признака).

$$F_3(x) = \frac{n_x}{n}, \text{ где } n_x = \sum_{x'_i < x} f_i, \text{ т.е. } n_x \text{ — сумма частот (накопленная}$$

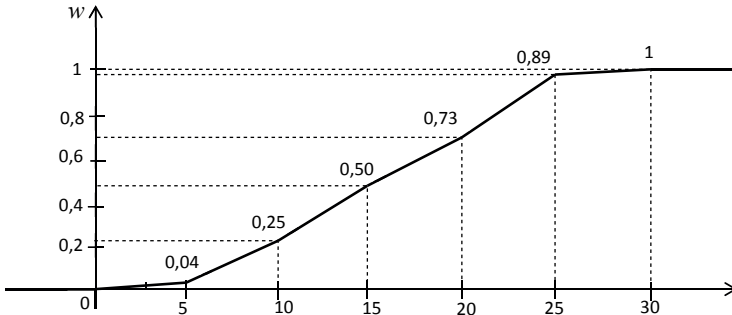
частота) тех вариантов  $x'_i$  (середин интервалов), которые меньше  $x$ .

Строим эмпирическую функцию распределения:



Здесь  $w$  — относительная частота того, что значение признака интервала, т. е.  $x'_i$ , меньше заданного значения  $x$ :  $w = F_{\Sigma}(x) = W(x'_i < x) = \frac{n_x}{n}$ .

12. Кумуляту строим также по значениям накопленных относительных частот, соединяя отрезками прямых соседние точки  $(x_i, F_{\Sigma}(x_i))$ , где  $x_i$  — конец  $i$ -го интервала:



Значение кумуляты в точке  $x$  дает приближенную оценку накопленной частоты тех значений уже самого признака (а не середин интервалов), рассматриваемого как непрерывная переменная, которые меньше заданного значения  $x$ . В нашем случае этим признаком является время, уделяемое студентами в день ведению домашнего хозяйства.

## *Задача № 4*

### **Выборочный метод в социологии**

**Цель:** отработка методов математического анализа количественных данных социологического опроса.

**Задание.** Провести анализ представленных выборочных данных по заданному алгоритму с целью определения показателей генеральной совокупности.

Алгоритм действий:

- 1) определить среднее значение показателя в генеральной совокупности, гарантируя результат с заданной надежностью;
- 2) определить долю элементов генеральной совокупности, обладающих заданным значением показателя, гарантируя результат с заданной надежностью;
- 3) найти необходимую численность выборки при определении среднего значения показателя, чтобы с заданной надежностью предельная ошибка выборки не превышала заданного значения;
- 4) найти необходимую численность выборки при определении доли элементов генеральной совокупности, обладающих заданным значением показателя, чтобы с заданной надежностью предельная ошибка выборки не превышала заданного значения.

### ***Пример выполнения работы\****

На предприятии в порядке случайной бесповторной выборки были опрошены 100 молодых сотрудников из 1000, в результате чего были получены следующие данные об их доходе за октябрь:

Месячный доход, руб.	600—1000	1000—1400	1400—1800	1800—2200
Число сотрудников	12	60	20	8

Требуется:

- 1) определить среднемесячный размер дохода у молодых сотрудников данного предприятия, гарантируя результат с вероятностью 0,997;



2) определить долю сотрудников предприятия, имеющих месячный доход 1400 руб. и выше, гарантируя результат с вероятностью 0,954;

3) найти необходимую численность выборки при определении среднего месячного дохода сотрудников предприятия, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 50 руб;

4) найти необходимую численность выборки при определении доли сотрудников с размером месячного дохода 1400 руб. и выше, чтобы с вероятностью 0,954 предельная ошибка не превышала 4 %.

### Решение

1. Генеральная средняя лежит в доверительном интервале:  $\bar{x} - \Delta_{\bar{x}} \leq \bar{X} \leq \bar{x} + \Delta_{\bar{x}}$ . Для ее определения нужно рассчитать выборочную среднюю:

$$\bar{x} = \frac{\sum_{i=1}^k x'_i f_i}{\sum_{i=1}^k f_i}.$$

$x_i$	$f_i$	$x'_i$	$x'_i f_i$	$(x'_i - \bar{x})$	$(x'_i - \bar{x})^2$	$(x'_i - \bar{x})^2 f_i$
600–1000	12	800	9600	– 496	246016	2952192
1000–1400	60	1200	72000	– 96	9216	552960
1400–1800	20	1600	32000	304	92416	1848320
1800–2200	8	2000	16000	704	495616	3964928
Итого ( $\Sigma$ )	100	–	129600	–	–	9318400

$$\bar{x} = \frac{129600}{100} = 1296.$$

2. Определить предельную ошибку выборки  $\Delta_{\bar{x}} = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$ .

Для этого нужно:

а) рассчитать дисперсию  $\sigma^2 = \frac{\sum_{i=1}^k (x'_i - \bar{x})^2 f_i}{\sum_{i=1}^k f_i}$  ;

$$\sigma^2 = \frac{9318400}{100} = 93184 ;$$

б) определить коэффициент доверия  $t$ , используя таблицу значений функций Лапласа, из которой видно, что для  $\gamma = 0,997$   $t = 3,00$ ;

Таблица 1

Значения функции Лапласа

$\gamma$	0,683	0,950	0,954	0,990	0,997
$t$	1,00	1,96	2,00	2,58	3,00

в) найти отношение  $\frac{n}{N} = \frac{100}{1000} = 0,10$  ;

г) подставить данные в формулу:

$$\Delta_{\bar{x}} = 3 \sqrt{\frac{93184}{100} (1 - 0,1)} = 86,878674 \approx 86,88 .$$

3. Зная выборочную среднюю и величину предельной ошибки выборки, можем определить пределы, в которые заключена генеральная средняя:

$$(1296 - 86,88 \leq \bar{X} \leq 1296 + 86,88) ,$$

$$(1209,12 \leq \bar{X} \leq 1382,88) .$$

Вывод: с вероятностью 0,997 мы можем утверждать, что среднемесячный размер дохода у молодых сотрудников данного предприятия не ниже 1209 руб. и не выше 1383 руб.

2. Определить долю сотрудников предприятия, имеющих месячный доход 1400 руб. и выше, гарантируя результат с вероятностью 0,954.

### Решение

Генеральная доля лежит в доверительном интервале  $w - \Delta_p \leq p \leq w + \Delta_p$ .

Для ее определения:

1) найдем долю сотрудников предприятия в выборочной совокупности, у которых месячный доход составляет 1400 руб. и выше:

$x_i$	$f_i$	$w_i$
До 1400	72	0,72
1400 и выше	28	0,28
Итого ( $\Sigma$ ):	100	1,00

То есть,  $w = 0,28$ .

2) найдем предельную ошибку выборки  $\Delta_p = t \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$ :

а)  $\gamma = 0,954$   $t = 2,00$  (см. табл. 1);

б) найдем отношение  $\frac{n}{N} = 0,10$ .

Подставим данные в формулу:

$$\Delta_p = 2 \sqrt{\frac{0,28(1-0,28)}{100}} (1-0,1) \approx 0,085;$$

3) зная выборочную долю и величину предельной ошибки выборки, можно определить пределы, в которых заключена генеральная доля:

$$(0,28 - 0,085 \leq p \leq 0,28 + 0,085),$$

$$(0,195 \leq p \leq 0,365).$$

Вывод: с вероятностью 0,954 мы можем утверждать, что доля сотрудников предприятия, имеющих месячный доход 1400 руб. и выше, не ниже 19,5 % и не выше 36,5 %.

3. Найти необходимую численность выборки при определении среднего месячного дохода сотрудников предприятия, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 50 руб.

***Решение***

Объем случайной бесповторной выборки для определения среднего значения количественного признака рассчитывается по формуле

$$n \geq \frac{t^2 \sigma^2 N}{t^2 \sigma^2 + N \Delta_{\bar{x}}^2}.$$

Нам задана допустимая предельная ошибка выборки, она составляет 50 руб. Коэффициент доверия мы можем найти по табл. 1:  $t = 2$ . Согласно предыдущему исследованию дисперсия данного признака составляет 93184. Эти данные мы можем использовать для определения объема выборки. Объем генеральной совокупности равен 1000 человек.

Подставим данные в формулу:

$$n \geq \frac{2^2 \cdot 93184 \cdot 1000}{2^2 \cdot 93184 + 1000 \cdot 50^2} \approx 129,75; \quad n = 130.$$

Ответ: необходимый объем выборки составляет 130 человек.

4. Найти необходимую численность выборки при определении доли сотрудников с размером месячного дохода 1400 руб. и выше, чтобы с вероятностью 0,954 предельная ошибка не превышала 4 %.

***Решение***

Объем случайной бесповторной выборки для определения доли единиц, обладающих данным значением признака, рассчитывается по формуле

$$n \geq \frac{t^2 w(1-w)N}{t^2 w(1-w) + N \Delta_p^2}.$$

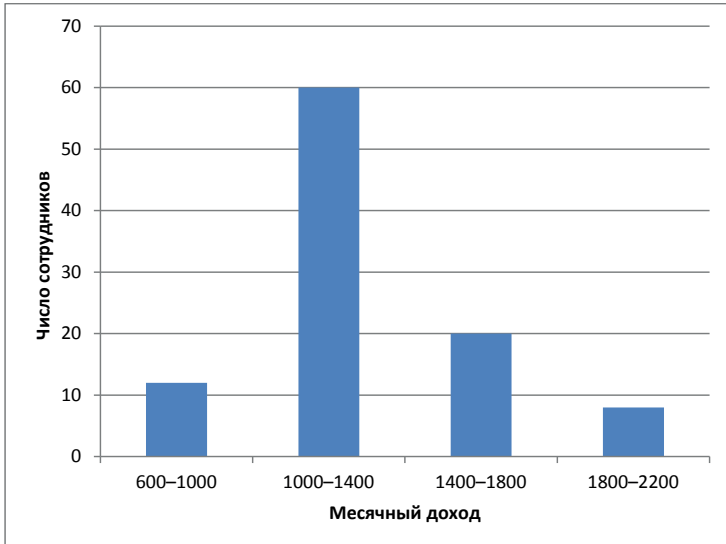
Нам задана допустимая предельная ошибка выборки (она составляет 4%, т.е. 0,04). Коэффициент доверия мы можем найти по табл. 1:  $t = 2$ . Согласно предыдущему исследованию доля рабочих с размером месячного дохода 1400 руб. и выше составляет 0,28. Эти данные мы можем использовать для определения объема выборки. Объем генеральной совокупности равен 1000 человек.

Подставим данные в формулу:

$$n \geq \frac{2^2 \cdot 0,28 \cdot (1 - 0,28) \cdot 1000}{2^2 \cdot 0,28 \cdot (1 - 0,28) + 1000 \cdot 0,04^2} \approx 335,11; \quad n = 336.$$

Ответ: необходимый объем выборки составляет 336 человек.

#### 5. Сравнительная диаграмма исследуемого признака.



### *Задача № 5*

#### **Анализ модели поведения социальной группы**

**Цель:** отработка методов математического анализа модели группового поведения.

**Задание.** Представлены данные о частоте того или иного поведения представителей определенных социальных групп. Провести сравнительный анализ поведения различных групп исследуемой совокупности.

**Порядок выполнения:**

- 1) определить условные индексы по группам и по массиву данных;
- 2) рассчитать предельную ошибку выборки для определения значимости различий между группами;
- 3) построить сравнительную диаграмму, найти моду и медиану признака для каждой группы и по массиву в целом;
- 4) провести социологический анализ полученных результатов.

#### ***Пример выполнения работы***

Ниже представлены данные о частоте просмотра телевизора студентами одного из вузов, живущими на свои заработки (самостоятельными) и находящимися на иждивении родственников (зависимыми). Проинтерпретируйте результаты.

Частота просмотра телевизора в зависимости от образа жизни\*  
(в абсолютных числах)

Частота просмотра	В целом по массиву	Образ жизни	
		Самостоя- тельные	Зависимые
Каждый день	148	48	100
4–5 дней в неделю	56	15	41
2–3 дня в неделю	76	29	47
1 день в неделю	19	10	9
Реже одного дня в неделю	42	25	17
Затрудняюсь ответить	29	10	19
Итого	370	137	233

Алгоритм решения:

- 1) рассчитать условный индекс для каждой группы и в целом по массиву по формуле средней арифметической взвешенной;
- 2) рассчитать предельную ошибку выборки для определения значимости различий между группами;
- 3) построить сравнительную диаграмму, найти моду и медиану признака для каждой группы и по массиву в целом;
- 4) представить результаты в текстовом виде;
- 5) сделать социологический вывод относительно полученных результатов.

### *Решение*

1. Средняя арифметическая взвешенная:

$$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^k f_i},$$

где  $m$  — число групп в массиве,  $x_i$  — признак группы, а  $k$  — число разрядов (градаций) основного признака. В рассматриваемом случае все  $x_i = 1$  (признак каждой группы атрибутивный), поэтому индексом будут проценты (отношение числа респондентов, ответивших утвердительно, к числу опрошенных, т. е. доля, выраженная в процентах).

2. Процентное деление совокупности по группам:

— ведущие самостоятельный образ жизни:  $\frac{137}{370} = 0,37 = 37\%$ .

— ведущие зависимый образ жизни:  $\frac{233}{370} = 0,63 = 63\%$ .

3. По массиву в целом:

— смотрят каждый день:

$$\bar{x} = \frac{100 + 48}{370} = 0,4000 = 40,00\%;$$

— смотрят 4–5 дней в неделю:

$$\bar{x} = \frac{15+41}{370} \approx 0,1514 = 15,14\%;$$

— смотрят 2–3 дня в неделю:

$$\bar{x} = \frac{29+47}{370} \approx 0,2054 = 20,54\%;$$

— смотрят 1 день в неделю:

$$\bar{x} = \frac{10+9}{370} \approx 0,0513 = 5,13\%;$$

— смотрят реже одного дня в неделю:

$$\bar{x} = \frac{25+17}{370} \approx 0,1135 = 11,35\%;$$

— затрудняются ответить:

$$\bar{x} = \frac{10+19}{370} \approx 0,0784 = 7,84\%.$$

4. По группе ведущих самостоятельный образ жизни:

— смотрят каждый день:

$$\bar{x} = \frac{48}{137} \approx 0,3503 = 35,03\%;$$

— смотрят 4–5 дней в неделю:

$$\bar{x} = \frac{15}{137} \approx 0,1095 = 10,95\%;$$

— смотрят 2–3 дня в неделю:

$$\bar{x} = \frac{29}{137} \approx 0,2117 = 21,17\%;$$



— смотрят 1 день в неделю:

$$\bar{x} = \frac{10}{137} \approx 0,0730 = 7,30\%;$$

— смотрят реже одного дня в неделю:

$$\bar{x} = \frac{25}{137} \approx 0,1825 = 18,25\%;$$

— затрудняются ответить:

$$\bar{x} = \frac{10}{137} \approx 0,0730 = 7,30\%.$$

5. По группе ведущих зависимый образ жизни:

— смотрят каждый день:

$$\bar{x} = \frac{100}{233} \approx 0,4292 = 42,92\%;$$

— смотрят 4–5 дней в неделю:

$$\bar{x} = \frac{41}{233} \approx 0,1760 = 17,60\%;$$

— смотрят 2–3 дня в неделю:

$$\bar{x} = \frac{47}{233} \approx 0,2017 = 20,17\%;$$

— смотрят 1 день в неделю:

$$\bar{x} = \frac{9}{233} \approx 0,0386 = 3,86\%;$$

— смотрят реже одного дня в неделю:

$$\bar{x} = \frac{17}{233} \approx 0,0730 = 7,30\%;$$

— затрудняются ответить:

$$\bar{x} = \frac{19}{233} \approx 0,0815 = 8,15\%.$$

6. Предельная ошибка выборки:

$$\Delta_w = t \sqrt{\frac{w(1-w)}{n}}, \text{ где } t \text{ — нормированное отклонение (коэффициент}$$

ент доверия): при вероятности (надежности)  $\gamma = \Phi(t) = 0,950$   $t = 1,96$ .

7. По массиву в целом:

смотрят каждый день:

$$w = \frac{148}{370} = 0,4000 = 0,4.$$

$$\Delta_1 = 1,96 \sqrt{\frac{0,4(1-0,4)}{370}} = 1,96 \sqrt{0,00065} \approx 0,0499 = 4,99\%.$$

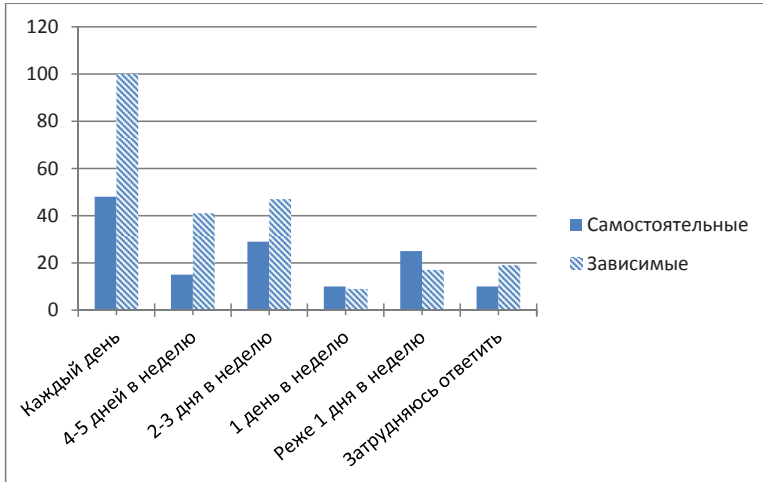
Для остальных показателей расчет ведется точно так же.

Аналогично рассчитываем предельную ошибку выборки и по группам ведущих самостоятельный и зависимый образ жизни.

Результаты вычислений представим в табличной форме.

Частота просмотра	В целом по массиву		Образ жизни			
	Ин- декс, %	Ошибка выбор- ки, %	Самостоятельные	Зависимые	Ин- декс, %	Ошибка выбор- ки, %
Каждый день	40,00	4,99	35,03	7,99	42,92	6,36
4–5 дней в неделю	15,14	3,65	10,95	5,23	17,60	4,89
2–3 дня в неделю	20,54	4,12	21,17	6,84	20,17	5,15
1 день в неделю	5,13	2,25	7,30	4,36	3,86	2,47
Реже одного дня в неделю	11,35	3,23	18,25	6,47	7,30	3,34
Затрудняюсь ответить	7,84	2,74	7,30	4,36	8,15	3,51
Итого	100	—	100	—	100	—

Эти же данные можно отразить и с помощью сравнительной диаграммы:



Предельная ошибка выборки используется для определения значимости различий между группами. Так, разница между количеством выборов альтернатив «Каждый день» и «4–5 дней в неделю» по всему массиву составляет  $40\% - 15,14\% = 34,86\%$ , а суммарная ошибка репрезентативности составляет  $4,99\% + 3,65\% = 8,64\%$ , что существенно меньше разницы между количеством выборов. Следовательно, можно утверждать, что в генеральной совокупности на самом деле число смотрящих телевизор каждый день значительно больше, чем тех, кто смотрит его 4–5 дней в неделю. Иная ситуация складывается при сравнении альтернатив «4–5 дней в неделю» и «2–3 дня в неделю». Разница между количеством выборов здесь составляет  $20,54\% - 15,14\% = 5,4\%$ . Суммарная ошибка репрезентативности равна  $4,12\% + 3,65\% = 7,77\%$ , что превышает разницу между количеством выборов. Таким образом, мы не можем быть уверены, что в действительности первая альтернатива выбирается чаще, чем вторая.

Аналогичный анализ может быть проведен при сравнении других пар альтернатив по всем группам.

### **Мода**

Модальным является признак, имеющий наибольшую частоту. Это признак первой группы:  $n_1 = 148$ .

Следовательно, значение модального признака — «Каждый день».

Для ведущих самостоятельный и ведущих зависимый образ жизни модальный признак остается тем же.

### **Медиана**

Медиана делит вариационный ряд на две равные по численности части.  $370/2 = 185$ . Элементы 185-й и 186-й попадают во вторую группу.

Следовательно, значение медианного признака — «4–5 раз в день».

Это значит, что половина респондентов смотрит телевизор не менее 4–5 раз в неделю, а половина — не более 4–5 раз в неделю или затрудняется с ответом. Если исключить затрудняющихся с ответом, то  $370 - 29 = 341$ , и серединой нового вариационного ряда будет 171-й элемент. Это элемент снова попадает во вторую группу, и значение медианного признака остается прежним — «4–5 раз в день», т. е. половина респондентов смотрит телевизор не менее 4–5 раз в неделю, а половина — не более 4–5 раз в неделю.

Для ведущих зависимый образ жизни (без затрудняющихся с ответом)  $n = 233 - 19 = 214$ , и серединой вариационного ряда будут 107-й и 108-й элементы. Они попадают во вторую группу. Значит, для этой группы вывод тот же.

Для ведущих самостоятельный образ жизни (без затрудняющихся с ответом)  $n = 137 - 10 = 127$ . и серединой вариационного ряда будет 64-й элемент, попадающий уже в третью группу. Значит, медианным признаком для них будет «2–3 раза в неделю»,

т.е. половина респондентов смотрит телевизор не менее 2–3 раз в неделю, а половина — не более 2–3 раз в неделю.

Итак, согласно полученным данным 40 % студентов (с ошибкой  $\pm 5\%$ ) смотрят телевизор ежедневно. Среди ведущих самостоятельный образ жизни таких оказалось 35 %, а среди ведущих зависимый образ жизни — 43 %.

Довольно сильно отличаются также доли тех, кто смотрит телевизор 4–5 раз в неделю. Таких среди ведущих самостоятельный образ жизни 11 %, а среди живущих на иждивении родственников — 18 %. Всего же только 15 % студентов смотрят телевизор 4–5 раз в неделю. Практически одинаков процент тех, кто смотрит телевизор 2–3 раза в неделю: 21 % среди самостоятельных и 20 % среди зависимых. К тому же, разность в один процент находится в пределах ошибок выборки. Разность снова становится существенной среди тех, кто смотрит телевизор реже одного раза в неделю: их среди ведущих самостоятельный образ жизни 18 %, а среди ведущих зависимый образ жизни — только 7 %. Эта разность больше суммарной ошибки выборки:  $6\% + 3\% = 9\%$ . Число затруднившихся ответить на вопрос практически одинаково: 7 % и 8 % соответственно.

Таким образом, наибольшая доля студентов, ведущих зависимый образ жизни, смотрит телевизор ежедневно. Число студентов, ведущих самостоятельный образ жизни и ежедневно смотрящих телевизор, значительно меньше. Это объясняется тем, что у живущих самостоятельно студентов намного меньше свободного времени. Этой же причиной объясняется и то, что процент тех, кто смотрит телевизор реже одного дня в неделю, среди ведущих самостоятельный образ жизни значительно выше, чем у тех, кто находится на иждивении родственников.

## *Задача № 6*

### **Сравнительный анализ распределения общественного мнения в разных социальных группах**

**Цель:** отработка математических методов анализа общественного мнения.

**Задание.** По представленным данным ранжированной оценки системы высказываний проанализировать распределение общественного мнения в разных социальных группах.

Порядок выполнения:

- 1) рассчитать условные индексы групп;
- 2) рассчитать ошибку выборки, оценить значимость различий между значениями индекса в разных группах;
- 3) представить результаты в графической форме;
- 4) сделать социологический вывод относительно полученных результатов

#### ***Пример выполнения работы\****

В табл. 2 представлены результаты опроса студентов по поводу их отношения к телевидению. Им был предложен вопрос:

«Охарактеризуйте ваше отношение к телевидению в целом.

Из пары предложенных противоположных высказываний выберите то, с которым вы согласны, и отметьте, насколько. Обведите кружком по одной цифре в каждой строке».

*Таблица 2*

#### Результаты опроса студентов

Телевидение показывает много захватывающих фильмов	3	2	1	0	1	2	3	На экране слишком много агрессивных фильмов и передач
Каждый может выбрать телепередачу по вкусу	3	2	1	0	1	2	3	Телевидение учитывает вкусы не всех зрителей

После просмотра телевизора я начинаю чувствовать себя подавленно и угнетенно	3	2	1	0	1	2	3	После просмотра телевизора я нахожусь в приподнятом настроении
Телевидение необходимо современному человеку	3	2	1	0	1	2	3	Просмотр телевизора — зря потраченное время
Большинство передач и художественных фильмов мне не интересны	3	2	1	0	1	2	3	Мне нравится почти все, что показывают по телевизору
Передачи и фильмы показывают реальную жизнь	3	2	1	0	1	2	3	По телевидению показывают искаженную жизнь

Результаты обработки данных представлены в табличной форме.

Выраженность признака	Пол		Пол		Пол	
	Мужской	Женский	Мужской	Женский	Мужской	Женский
	Телевидение учитывает вкусы всех телезрителей		Телевидение необходимо современному человеку		Просмотр телевизора улучшает мое настроение	
1	40	55	52	45	16	8
2	37	61	32	45	41	31
3	28	20	28	41	27	41
4	26	23	31	35	81	90
5	11	9	17	13	5	11
6	23	14	13	5	7	5
7	14	7	6	5	2	3
Всего	179	189	179	189	179	189

Выра- жен- ность признака	Телевидение по- зволяет видеть со- бытия в динамике		Телевидение по- казывает жизнь такой, какая она есть		Мне нравится почти все, что по- казывают по теле- визору	
1	20	8	6	8	4	2
2	25	29	13	16	10	10
3	31	36	27	35	20	26
4	39	32	58	68	43	50
5	30	19	33	25	40	29
6	22	26	26	21	33	49
7	12	39	16	16	29	23
Всего	179	189	179	189	179	189

Ранжирование велось по нарастанию выраженности признака, т. е. минимальная выраженность оценивалась значением 1, а максимальная — значением 7.

Необходимо проанализировать полученные результаты.

Алгоритм решения:

- 1) рассчитать индексы по формуле средней арифметической взвешенной для каждой из групп;
- 2) рассчитать ошибку выборки, оценить значимость различий между значениями индекса в разных группах;
- 3) представить результаты в графической форме, располагая позиции в порядке возрастания/убывания различия между ними;
- 4) описать результаты в текстовой форме;
- 5) сделать социологический вывод относительно полученных результатов.

### ***Решение***

1. Для вычисления индексов по формуле средней арифметической взвешенной заполним расчетную таблицу.



Выраженность признака	Пол		Пол		Пол		Пол		Пол		Пол		Пол		Пол	
	Муж-ской	Жен-ской	Муж-ской	Жен-ской	Муж-ской	Жен-ской	Муж-ской	Жен-ской	Муж-ской	Жен-ской	Муж-ской	Жен-ской	Муж-ской	Жен-ской	Муж-ской	Жен-ской
	Телевидение учитывает вкусы всех телезрителей		Телевидение необходимо современному человеку		Просмотр телевизора улучшает мое настроение		Телевидение позволяет видеть события в динамике		Телевидение показывает жизнь такой, какая она есть		Телевидение		Мне нравится почти все, что показывают		по телевизору	
$x_i$	$f_i$	$xf_i$	$f_i$	$xf_i$	$f_i$	$xf_i$	$f_i$	$xf_i$	$f_i$	$xf_i$	$f_i$	$xf_i$	$f_i$	$xf_i$	$f_i$	$xf_i$
1	40	40	55	55	52	52	45	45	16	16	8	8	20	20	8	8
2	37	74	61	122	32	64	45	90	41	82	31	62	25	50	29	58
3	28	84	20	60	28	84	41	123	27	81	41	123	31	93	36	108
4	26	104	23	92	31	124	35	140	81	324	90	360	39	156	32	128
5	11	55	9	45	17	85	13	65	5	25	11	55	30	150	19	95
6	23	138	14	84	13	78	5	30	7	42	5	30	22	132	26	156
7	14	98	7	49	6	42	5	35	2	14	3	21	12	84	39	273
Всего	179	593	189	507	179	529	189	528	179	584	189	659	179	685	189	826
го													179	778	189	780
													4	4	4	4
													10	20	10	20
													20	60	26	78
													43	172	50	200
													40	200	29	145
													33	198	49	294
													29	203	23	161
													179	857	189	900

Формула для расчета средней арифметической взвешенной:  $\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i}$ . Например,

$$\bar{x}_{1M} = \frac{593}{179} = 3,31, \quad \bar{x}_{1Ж} = \frac{507}{189} = 2,68.$$

2. Для подсчета дисперсии и предельной ошибки выборки составим еще одну таблицу.

Выраженность признака	Пол		Пол		Пол		Пол		Пол		Пол	
	Муж-ский	Жен-ский	Муж-ский	Жен-ский	Муж-ский	Жен-ский	Муж-ский	Жен-ский	Муж-ский	Жен-ский	Муж-ский	Жен-ский
Телевидение учитывает вкусы всех телезрителей	Телевидение необходимо современному человеку	Просмотр теле-визора улучшает мое настроение	Телевидение позволяет видеть события в динамике	Телевидение по-казывает жизнь такой, какая она есть	Мне нравится почти все, что показывают по телевизору							
$x_i$	$x_i f_i$	$x_i^2 f_i$	$x_i f_i$	$x_i^2 f_i$	$x_i f_i$	$x_i^2 f_i$	$x_i f_i$	$x_i^2 f_i$	$x_i f_i$	$x_i^2 f_i$	$x_i f_i$	$x_i^2 f_i$
1	40	55	52	45	16	8	20	20	8	6	4	2
2	74	148	122	244	64	128	90	180	82	164	62	124
3	84	252	60	180	84	252	123	369	81	243	123	369
4	104	416	92	368	124	496	140	560	324	1296	360	1440
5	55	275	45	225	85	425	65	325	25	125	55	275
6	138	828	84	504	78	468	30	180	42	252	30	180
7	98	686	49	343	42	294	35	245	14	98	21	147
Всего	593	2645	507	1919	529	2115	528	1904	584	2194	659	2543
го												

Формулы для расчета дисперсии:  $\sigma^2 = \overline{x^2} - (\overline{x})^2$ ,  $\overline{x^2} = \frac{\sum_{i=1}^k x_i^2 f_i}{\sum_{i=1}^k f_i}$ . Ошибка выборки

$$\Delta = t \sqrt{\frac{\sigma^2}{n}}, t = 1,96 \text{ при } \gamma = 0,950.$$

$$n = \sum_{i=1}^k f_i. \text{ Например, } \sigma_{\text{лм}}^2 = \frac{2645}{179} - (3,31)^2 = 3,82, \Delta_{\text{лм}} = 1,96 \sqrt{\frac{3,82}{179}} \approx 0,29. \text{ Остальные значения подсчитаем аналогично.}$$

3. Используя данные расчетных таблиц и приведенные формулы, найдем условные индексы, дисперсии и предельные ошибки выборки по каждой группе. Результаты отразим в табличной форме.

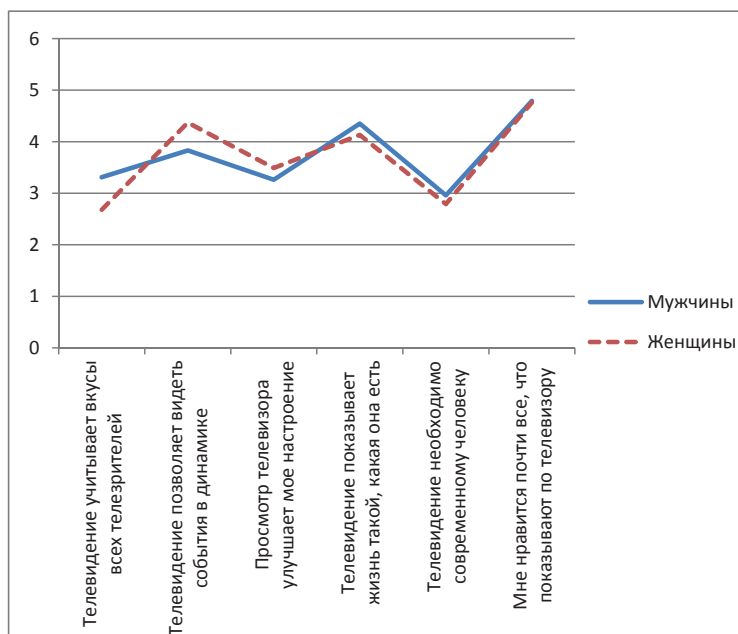
Пол		Пол		Пол		Пол		Пол		Пол	
Муж- ской	Жен- ский	Муж- ской	Жен- ский	Муж- ской	Жен- ский	Муж- ской	Жен- ский	Муж- ской	Жен- ский	Муж- ской	Жен- ский
Телевидение учитывает вкусы всех телезрителей		Телевидение необходимо современному человеку		Просмотр теле-визора улучшает мое настроение		Телевидение позволяет видеть события в динамике		Телевидение показывает жизнь такой, какая она есть		Мне нравится почти все, что показывают по телевизору	
3,31	2,68	2,96	2,79	3,26	3,49	3,83	4,37	4,35	4,13	4,79	4,76

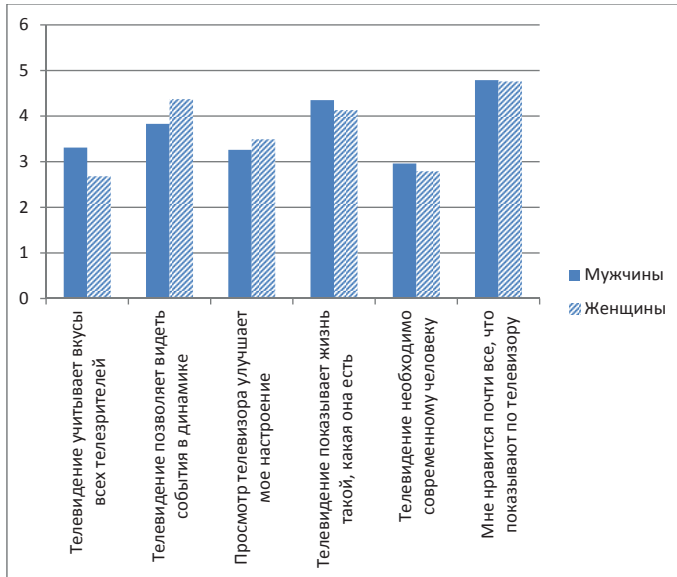
Представим результаты, располагая позиции в порядке убывания различия между ними:

$$\delta = \left| \bar{X}_{\text{муж}} - \bar{X}_{\text{жен}} \right|.$$

Признак	Мужчины		Женщины		δ
	$\bar{x}$	$\sigma^2$	$\bar{x}$	$\sigma^2$	
Телевидение учитывает вкусы всех телезрителей	3,31	3,82	2,68	2,97	0,25 0,63
Телевидение позволяет видеть события в динамике	3,83	2,95	0,25 4,37	3,56	0,27 0,54
Просмотр телевизора улучшает мое настроение	3,26	1,63	0,19 3,49	1,27	0,16 0,23
Телевидение показывает жизнь такой, какая она есть	4,35	2,16	0,22 4,13	2,20	0,21 0,22
Телевидение необходимо современному человеку	2,96	3,05	0,26 2,79	2,29	0,22 0,17
Мне нравится почти все, что показывают по телевизору	4,79	2,31	0,22 4,76	2,17	0,21 0,03

4. По первым двум значениям признака в составленной таблице разница значений индексов  $\bar{x}$  у мужчин и женщин больше суммарной ошибки репрезентативности ( $\delta_1 = 0,63$ ;  $\Delta_{1м} + \Delta_{1ж} = 0,29 + 0,25 = 0,54$ ;  $\delta_2 = 0,54$ ;  $\Delta_{2м} + \Delta_{2ж} = 0,25 + 0,27 = 0,52$ ), а по остальным значениям признака она меньше. Это значит, что по первым двум признакам оценки мужчин и женщин отличаются в генеральной совокупности, а отличия в индексах других групп могут быть вызваны ошибками выборки, и с уверенностью говорить о различии оценок нельзя. Представим результаты в графической форме.





По степени выраженности признака альтернативы располагаются следующим образом.

Признак	Мужчины Индекс	Признак	Женщины Индекс
Мне нравится почти все, что показывают по телевизору	4,79	Мне нравится почти все, что показывают по телевизору	4,76
Телевидение показывает жизнь такой, какая она есть	4,35	Телевидение позволяет видеть события в динамике	4,37
Телевидение позволяет видеть события в динамике	3,83	Телевидение показывает жизнь такой, какая она есть	4,13
Телевидение учитывает вкусы всех телезрителей	3,31	Просмотр телевизора улучшает мое настроение	3,49

Признак	Мужчины	Признак	Женщины
	Индекс		Индекс
Просмотр телевизора улучшает мое настроение	3,26	Телевидение необходимо современному человеку	2,79
Телевидение необходимо современному человеку	2,96	Телевидение учитывает вкусы всех телезрителей	2,68

5. Наибольшее различие во взглядах мужчин и женщин наблюдается по поводу высказывания «Телевидение учитывает вкусы всех телезрителей», причем больший скептицизм по этому поводу выражают женщины. При этом оценка справедливости высказывания у тех и других умеренно негативная: индекс у мужчин — 3,31, а у женщин — 2,68 (из максимального 7).

Есть различие во взглядах и по поводу высказывания «Телевидение позволяет видеть события в динамике», причем у мужчин отношение к нему скорее отрицательное, чем положительное (индекс — 3,83, меньше, чем 4) а у женщин скорее положительное, чем отрицательное (индекс — 4,13, больше, чем 4). Различие в остальных показателях лежит в пределах ошибки наблюдений (ошибки выборки).

В положительную сторону склоняются мнения мужчин по поводу высказываний «Мне нравится почти все, что показывают по телевизору» и «Телевидение показывает жизнь такой, какая она есть» (индексы выше, чем 4).

Мнения женщин склоняются в положительную сторону относительно высказываний «Мне нравится почти все, что показывают по телевизору», «Телевидение позволяет видеть события в динамике» и «Телевидение показывает жизнь такой, какая она есть» (индексы выше, чем 4).

Наиболее скептическое отношение у мужчин к высказыванию «Телевидение необходимо современному человеку» (индекс — 2,96), а у женщин к высказыванию «Телевидение учитывает вкусы всех телезрителей» (индекс — 2,68).

## Задача № 7

### Системы ценностей в социальных группах

**Цель:** отработать методы математического анализа систем ценностей.

**Задание.** По данным ранжирования ценностей совокупностью респондентов построить ранжированный ряд ценностей по группе в целом.

Порядок выполнения:

- 1) рассчитать условный индекс для каждой ценности;
- 2) проранжировать ценности по величине индекса;
- 3) рассчитать предельную ошибку выборки для определения значимости различий между ценностями;
- 4) представить ранжированный ряд в текстовом виде;
- 5) сделать социологический вывод относительно полученных результатов.

### Пример выполнения работы

Ниже представлены результаты ранжирования ценностей по значимости студентами одного из вузов Урала. Выстройте ранжированный ряд ценностей в целом по массиву. Проинтерпретируйте результаты.

Результаты ранжирования ценностей по значимости\*  
(в абсолютных числах)

Ранг	День- ги	Се- мья	Здо- ро- вье	Инте- ресная работа	Друж- ба	Кра- сота	Лю- бовь	Ста- тус	Само- реа- лиза- ция	Рели- гия
1	31	71	101	10	35	2	53	6	21	4
2	58	70	57	17	36	0	53	8	10	4
3	55	52	58	32	35	5	44	6	11	4
4	43	35	36	42	45	2	43	8	21	4
5	35	22	20	62	54	12	30	17	14	4
6	21	20	19	41	46	19	29	17	17	5
7	34	18	20	29	22	13	25	28	24	2

Ранг	День- ги	Се- мья	Здо- ро- вье	Инте- ресная работа	Друж- ба	Кра- сота	Лю- бовь	Ста- тус	Само- реа- лиза- ция	Ре- ли- гия
8	16	16	14	39	24	17	20	19	28	6
9	10	13	9	23	20	16	24	28	29	7
10	12	11	10	16	14	16	11	44	17	8
11	11	6	7	16	13	16	3	41	41	6
12	9	6	4	13	6	26	12	36	32	8
13	8	7	2	6	9	43	9	27	40	14
14	11	9	5	11	4	43	4	29	29	14
15	7	7	4	6	3	61	5	25	20	21
16	5	3	4	5	3	48	3	15	11	62
17	4	4	0	2	1	31	2	16	5	197
Ито- го	370	370	370	370	370	370	370	370	370	370

### Решение

1. Рассчитать условный индекс для каждой ценности по формуле средней арифметической взвешенной.

Формула для расчета средней арифметической взвешенной

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i}.$$

Расчеты проведем по каждой группе:

$$\text{— «Деньги»: } \bar{x} = \frac{1 \cdot 31 + 2 \cdot 58 + \dots + 16 \cdot 5 + 17 \cdot 4}{370} = \frac{2101}{370} = 5,68;$$

$$\text{— «Семья»: } \bar{x} = \frac{1 \cdot 71 + 2 \cdot 70 + \dots + 16 \cdot 3 + 17 \cdot 4}{370} = \frac{1794}{370} = 4,85;$$

$$\text{— «Здоровье»: } \bar{x} = \frac{1 \cdot 101 + 2 \cdot 57 + \dots + 16 \cdot 4 + 17 \cdot 0}{370} = \frac{1525}{370} = 4,12;$$



$$\text{— «Интересная работа»: } \bar{x} = \frac{1 \cdot 10 + 2 \cdot 17 + \dots + 16 \cdot 5 + 17 \cdot 2}{370} = \frac{2514}{370} = 6,79;$$

$$\text{— «Дружба»: } \bar{x} = \frac{1 \cdot 35 + 2 \cdot 36 + \dots + 16 \cdot 3 + 17 \cdot 1}{370} = \frac{2102}{370} = 5,68;$$

$$\text{— «Красота»: } \bar{x} = \frac{1 \cdot 2 + 2 \cdot 0 + \dots + 16 \cdot 48 + 17 \cdot 31}{370} = \frac{4589}{370} = 12,4;$$

$$\text{— «Любовь»: } \bar{x} = \frac{1 \cdot 53 + 2 \cdot 53 + \dots + 16 \cdot 3 + 17 \cdot 2}{370} = \frac{1955}{370} = 5,28;$$

$$\text{— «Статус»: } \bar{x} = \frac{1 \cdot 6 + 2 \cdot 8 + \dots + 16 \cdot 15 + 17 \cdot 16}{370} = \frac{3826}{370} = 10,34;$$

$$\text{— «Самореализация»: } \bar{x} = \frac{1 \cdot 21 + 2 \cdot 10 + \dots + 16 \cdot 11 + 17 \cdot 5}{370} = \frac{3475}{370} = 9,39;$$

$$\text{— «Религия»: } \bar{x} = \frac{1 \cdot 4 + 2 \cdot 4 + \dots + 16 \cdot 62 + 17 \cdot 197}{370} = \frac{5491}{370} = 14,84.$$

2. Рассчитать предельную ошибку выборки для определения значимости различий между ценностями.

$$\text{Формулы для расчета дисперсии: } \sigma^2 = \overline{x^2} - (\bar{x})^2, \quad \overline{x^2} = \frac{\sum_{i=1}^k x_i^2 f_i}{\sum_{i=1}^k f_i}.$$

$$\text{Ошибка выборки } \Delta = t \sqrt{\frac{\sigma^2}{n}}, \quad t = 1,96 \text{ при } \gamma = 0,950.$$

Для ценности «Деньги»:

$$\overline{x^2} = \frac{17923}{370} = 48,44; \quad \sigma^2 = 48,44 - (5,68)^2 = 16,20;$$

$$\Delta = 1,96 \sqrt{\frac{16,20}{370}} = 1,96 \sqrt{0,043784} = 0,410122 \approx 0,41.$$

Для ценности «Семья»:

$$\overline{x^2} = \frac{14744}{370} = 39,85; \sigma^2 = 39,85 - (4,85)^2 = 16,34;$$

$$\Delta = 1,96 \sqrt{\frac{16,34}{370}} = 1,96 \sqrt{0,044162} = 0,41189 \approx 0,41$$

Для ценности «Здоровье»:

$$\overline{x^2} = \frac{10881}{370} = 29,41; \sigma^2 = 29,41 - (4,12)^2 = 12,42;$$

$$\Delta = 1,96 \sqrt{\frac{12,42}{370}} = 1,96 \sqrt{0,033568} = 0,3591 \approx 0,36.$$

Для ценности «Интересная работа»:

$$\overline{x^2} = \frac{21630}{370} = 58,46; \sigma^2 = 58,46 - (6,79)^2 = 12,29;$$

$$\Delta = 1,96 \sqrt{\frac{12,29}{370}} = 1,96 \sqrt{0,033216} = 0,357216 \approx 0,36.$$

Для ценности «Дружба»:

$$\overline{x^2} = \frac{16328}{370} = 44,13; \sigma^2 = 44,13 - (5,68)^2 = 11,86;$$

$$\Delta = 1,96 \sqrt{\frac{11,86}{370}} = 1,96 \sqrt{0,032054} = 0,350911 \approx 0,35.$$

Для ценности «Красота»:

$$\overline{x^2} = \frac{62031}{370} = 167,65; \sigma^2 = 167,65 - (12,40)^2 = 13,82;$$

$$\Delta = 1,96\sqrt{\frac{13,82}{370}} = 1,96\sqrt{0,037351} = 0,378799 \approx 0,38.$$

Для ценности «Любовь»:

$$\overline{x^2} = \frac{15559}{370} = 42,05; \sigma^2 = 42,05 - (5,28)^2 = 14,13;$$

$$\Delta = 1,96\sqrt{\frac{14,13}{370}} = 1,96\sqrt{0,038189} = 0,383024 \approx 0,38.$$

Для ценности «Статус»:

$$\overline{x^2} = \frac{44994}{370} = 121,61; \sigma^2 = 121,61 - (10,34)^2 = 14,68;$$

$$\Delta = 1,96\sqrt{\frac{14,68}{370}} = 1,96\sqrt{0,039676} = 0,390408 \approx 0,39.$$

Для ценности «Самореализация»:

$$\overline{x^2} = \frac{39249}{370} = 106,08; \sigma^2 = 106,08 - (9,39)^2 = 17,87;$$

$$\Delta = 1,96\sqrt{\frac{17,87}{370}} = 1,96\sqrt{0,048297} = 0,430742 \approx 0,43.$$

Для ценности «Религия»:

$$\overline{x^2} = \frac{86767}{370} = 234,51; \sigma^2 = 234,51 - (14,84)^2 = 14,26;$$

$$\Delta = 1,96\sqrt{\frac{14,26}{370}} = 1,96\sqrt{0,038541} = 0,384782 \approx 0,38.$$

Данные расчета взвешенной средней, дисперсии и предельной ошибки представлены в табличной форме.

Ранг	Деньги		Семья		Здоровье		Интерес- ная работа		Дружба	
$x$	$xf$	$x^2f$	$xf$	$x^2f$	$xf$	$x^2f$	$xf$	$x^2f$	$xf$	$x^2f$
1	31	31	71	71	101	101	10	10	35	35
2	116	232	140	280	114	228	34	68	72	144
3	165	495	156	468	174	522	96	288	105	315
4	172	688	140	560	144	576	168	672	180	720
5	175	875	110	550	100	500	310	1550	270	1350
6	126	756	120	720	114	684	246	1476	276	1656
7	238	1666	126	882	140	980	203	1421	154	1078
8	128	1024	128	1024	112	896	312	2496	192	1536
9	90	810	117	1053	81	729	207	1863	180	1620
10	120	1200	110	1100	100	1000	160	1600	140	1400
11	121	1331	66	726	77	847	176	1936	143	1573
12	108	1296	72	864	48	576	156	1872	72	864
13	104	1352	91	1183	26	338	78	1014	117	1521
14	154	2156	126	1764	70	980	154	2156	56	784
15	105	1575	105	1575	60	900	90	1350	45	675
16	80	1280	48	768	64	1024	80	1280	48	768
17	68	1156	68	1156	0	0	34	578	17	289
Итого	2101	17923	1794	14744	1525	10881	2514	21630	2102	16328
$\bar{x}$	5,68		4,85		4,12		6,79		5,68	
$\overline{x^2}$	48,44		39,85		29,41		58,46		44,13	
$\sigma^2$	16,20		16,34		12,42		12,29		11,86	
$\Delta$	0,41		0,41		0,36		0,36		0,35	

Ранг	Красота		Любовь		Статус		Самореализация		Религия	
$x$	$xf$	$x^2f$	$xf$	$x^2f$	$xf$	$x^2f$	$xf$	$x^2f$	$xf$	$x^2f$
1	2	2	53	53	6	6	21	21	4	4
2	0	0	106	212	16	32	20	40	8	16
3	15	45	132	396	18	54	33	99	12	36
4	8	32	172	688	32	128	84	336	16	64
5	60	300	150	750	85	425	70	350	20	100
6	114	684	174	1044	102	612	102	612	30	180
7	91	637	175	1225	196	1372	168	1176	14	98
8	136	1088	160	1280	152	1216	224	1792	48	384
9	144	1296	216	1944	252	2268	261	2349	63	567
10	160	1600	110	1100	440	4400	170	1700	80	800
11	176	1936	33	363	451	4961	451	4961	66	726
12	312	3744	144	1728	432	5184	384	4608	96	1152
13	559	7267	117	1521	351	4563	520	6760	182	2366
14	602	8428	56	784	406	5684	406	5684	196	2744
15	915	13725	75	1125	375	5625	300	4500	315	4725
16	768	12288	48	768	240	3840	176	2816	992	15872
17	527	8959	34	578	272	4624	85	1445	3349	56933
Итого	4589	62031	1955	15559	3826	44994	3475	39249	5491	86767
$\bar{x}$	12,40		5,28		10,34		9,39		14,8	
$\overline{x^2}$	167,65		42,05		121,61		106,08		234,51	
$\sigma^2$	13,82		14,13		14,68		17,87		14,26	
$\Delta$	0,38		0,38		0,39		0,43		0,38	

### 3. Проранжируем ценности по величине индекса.

Величина индекса	Ошибка выборки	Ранг	Ценность
4,12	0,36	1	Здоровье
4,85	0,41	2	Семья

Величина индекса	Ошибка выборки	Ранг	Ценность
5,28	0,38	3	Любовь
5,68	0,41	4–5	Деньги
5,68	0,35	4–5	Дружба
6,79	0,36	6	Интересная работа
9,39	0,43	7	Самореализация
10,34	0,39	8	Статус
12,40	0,38	9	Красота
14,84	0,38	10	Религия

4. Таким образом, на основании полученных результатов можно сделать вывод о том, что основная часть молодежи главными ценностями считает здоровье, семью и любовь. Последней по степени значимости является религия. Это говорит о том, что религия как духовная ценность мало интересует современную российскую молодежь.

Пары ценностей	Разность величин индексов	Суммарная ошибка выборки
Здоровье—семья	0,73	0,77
Семья—любовь	0,43	0,79
Здоровье—любовь	1,16	0,74
Любовь—деньги	0,40	0,79
Деньги—дружба	0,00	0,76
Дружба—интересная работа	1,11	0,71
Интересная работа—самореализация	2,60	0,79
Самореализация—статус	0,95	0,82
Статус—красота	2,06	0,77
Красота—религия	2,44	0,76

Различие величин индексов в парах здоровье—семья, семья—любовь, любовь—деньги, деньги—дружба меньше суммарной ошибки выборки (репрезентативности). Это значит, что в генеральной совокупности эти ценности близки по рангу и могут поменяться местами. Но уже при сравнении пары здоровье—любовь разность величин индексов ( $5,28 - 4,12 = 1,16$ ) больше суммарной ошибки выборки ( $0,36 + 0,38 = 0,74$ ). Это говорит о том, что здоровье как ценность ставится выше, чем любовь.

5. На основании такого сравнительного анализа можно сделать вывод, что рассматриваемые ценности довольно резко делятся на три группы. Первую группу составляют здоровье, семья и любовь. Это наиболее важные и близкие по рангу ценности, где соседние по рангу ценности могут в генеральной совокупности поменяться местами. Вторую группу составляют деньги и дружба. Это практически неразличимые по рангу ценности, т.е. нет возможности определить, какая из них является предпочтительной. Но их ранг ниже рангов ценностей первой группы. В третью группу (с рангами ниже рангов первой и второй группы) входят интересная работа, самореализация, статус, красота и религия. Ранги этих ценностей четко различаются, т.к. разность величин индексов значительно выше суммарной ошибки выборки (ошибки репрезентативности).

## *Задача № 8*

### **Статистическая оценка социологических гипотез**

**Цель:** отработать математические методы оценки и проверки гипотез.

**Задание.** По приведенным материалам социологического опроса проверить выдвинутую гипотезу.

Порядок выполнения:

- 1) выдвинуть нулевую и альтернативную гипотезы;
- 2) определить тип критической области;
- 3) выбрать необходимый критерий проверки (статистику);
- 4) найти критические точки с необходимым уровнем значимости;
- 5) рассчитать эмпирическое значение критерия;
- 6) сравнить эмпирическое значение критерия с критическими точками;
- 7) сделать соответствующий вывод по поводу проверки гипотезы;
- 8) проверить коррелированность ранжированных рядов.

### ***Пример выполнения работы***

С помощью данных опроса проверить гипотезу: ценностное сознание меняется с возрастом.

Предпочитаемые ценности в зависимости от возраста\*

Ценности	Моло- дежь	Взро- слые	Моло- дежь	Взро- слые
	Частоты		Процент от числа ответивших	
Счастливая семейная жизнь	35	123	72,92 %	80,92 %
Здоровье	35	113	72,92 %	74,34 %
Материально обеспеченная жизнь	31	142	64,58 %	93,42 %
Интересная работа	27	56	56,25 %	36,84 %



Ценности	Моло- дежь	Взро- слые	Моло- дежь	Взро- слые
	Частоты		Процент от числа ответивших	
Любовь	25	11	52,08 %	7,24 %
Наличие хороших и верных друзей	21	47	43,75 %	30,92 %
Карьера	11	6	22,92 %	3,95 %
Удовольствия	10	16	20,83 %	10,53 %
Общественное признание	7	52	14,58 %	34,21 %
Познание нового	7	20	14,58 %	13,16 %
Активная деятельная жизнь	4	74	8,33 %	48,68 %
Возможность творчества	2	5	4,17 %	3,29 %

Количество ответивших: молодежь — 48, взрослые — 152.

### Решение

Нужно проверить, не противоречат ли данные исследования гипотезе о том, что есть зависимость между ценностным сознанием и возрастом.

Зависимая переменная — значимые ценности. Это поливариантная переменная, представляющая собой набор номинальных переменных, которые можно упорядочить по частоте выбора. Независимая переменная — возраст респондента. Это порядковая переменная. Исследованы две группы респондентов, т.е. взяты два значения независимой переменной: «молодежь» и «взрослые». Полученные ряды нужно упорядочить по частоте выбора первой группы респондентов (в нашем случае — «молодежь»). Это уже сделано в приведенной таблице. Процент от числа ответивших отражает относительную частоту выбора, т.е. долю. Проценты также уже подсчитаны.

1. Выдвинем нулевую ( $H_0$ ) и альтернативную ( $H_1$ ) гипотезы.

Нулевая гипотеза ( $H_0$ ): ценностное сознание не зависит от воз-

раста, другими словами, различие в долях выбора ( $\Delta$ ) имеет случайный характер, оно не значимо ( $\Delta = 0$ ). Альтернативная гипотеза ( $H_1$ ): ценностное сознание зависит от возраста, различие в долях выбора не случайно ( $\Delta \neq 0$ ).

2. Критическая область двусторонняя (т. к.  $H_0: \Delta = 0$ , а  $H_1: \Delta \neq 0$ ).

Односторонняя критическая область будет в случае  $H_0: \Delta = 0$ , а  $H_1: \Delta > 0$  (правосторонняя) и  $H_0: \Delta = 0$ , а  $H_1: \Delta < 0$  (левосторонняя).

3. Оценим статистическую значимость различий между долями выборов ценностей в группах молодежи и взрослых. Воспользуемся  $t$ -критерием Стьюдента для сравнения долей:

$$t = \frac{w_1 - w_2}{\sqrt{\frac{w_1(1-w_1)}{n_1} + \frac{w_2(1-w_2)}{n_2}}},$$

где  $w_1$  — доля в первой группе,  $w_2$  — доля во второй группе,  $n_1$  — объем первой группы,  $n_2$  — объем второй группы.

4. Из стандартного набора значений надежности проверки (0,9; 0,95; 0,99; 0,999) выберем значение 0,99, т. е. 99 % надежность.

Для выбранного значения надежности  $\gamma = 0,99$  найдем уровень значимости (вероятность ошибки первого рода)  $\alpha = 1 - \gamma = 0,01$  и число степеней свободы  $k = n_1 + n_2 - 2$  (в нашем случае  $k = 48 + 152 - 2 = 198$ ). По табл. П. 1 (см. приложение) критических точек распределения Стьюдента для двусторонней области найдем  $t_{кр} = 2,576$  (для выборок большого объема, когда число степеней свободы больше 120, степень свободы распределения Стьюдента считают равной  $\infty$ ).

5. Среди показателей есть ценности, значительно отличающиеся по долям выбора. Например, среди взрослых значимость ценности «активная жизнь» отмечают почти 49 %, а среди молодежи — только 8 %. Проверим, не являются ли такие отклонения случайными.

Рассчитаем эмпирическое значение  $t$ -критерия Стьюдента:

$$t_{\text{эмп}} = \frac{w_1 - w_2}{\sqrt{\frac{w_1(1-w_1)}{n_1} + \frac{w_2(1-w_2)}{n_2}}} = \frac{0,0833 - 0,4868}{\sqrt{\frac{0,0833(1-0,0833)}{48} + \frac{0,4868(1-0,4868)}{152}}} = -7,095$$

6. Проверяем гипотезу: отклонения относительных частот случайны ( $H_0$ ). Если  $|t_{\text{эмп}}| > t_{\text{кр}}$ , то нулевая гипотеза отвергается. Это значит, что результаты наблюдения (с вероятностью ошибки, равной уровню значимости  $\alpha$ ) не подтверждают гипотезу. Тогда принимается альтернативная гипотеза. Если же  $|t_{\text{эмп}}| < t_{\text{кр}}$ , то результаты наблюдения не противоречат выдвинутой нулевой гипотезе и, следовательно, нет причин ее отвергать. Она принимается с надежностью  $\gamma$ , или, по-другому, с вероятностью ошибки  $\alpha$ .

7. В нашем случае  $t_{\text{кр}} = 2,576$ , а  $|t_{\text{эмп}}| = 7,095$ . Следовательно,  $|t_{\text{эмп}}| > t_{\text{кр}}$ , и значит, нулевая гипотеза должна быть отвергнута. Таким образом, выявленное различие является значимым (не случайным).

Аналогичное исследование можно провести и по другим ценностям. Результаты представлены в табличной форме.

Ценности	Моло- дежь	Взро- слые	$t_{\text{кр}} = 2,576$	
	Относитель- ная частота (доля) выбора		$t_{\text{эмп}}$	Различие
Счастливая семейная жизнь	0,7292	0,8092	-1,11698	Незначимо
Здоровье	0,7292	0,7434	-0,1938	Незначимо
Материально обеспеченная жизнь	0,6458	0,9342	-4,01102	Значимо
Интересная работа	0,5625	0,3684	2,378821	Незначимо
Любовь	0,5208	0,0724	5,970102	Значимо
Наличие хороших и верных друзей	0,4375	0,3092	1,58744	Незначимо
Карьера	0,2292	0,0395	3,025948	Значимо
Удовольствия	0,2083	0,1053	1,617398	Незначимо

Ценности	Моло- дежь	Взро- слые	$t_{кр} = 2, 576$	
	Относитель- ная частота (доля) выбора		$t_{эмп}$	Различие
Общественное признание	0,1458	0,3421	−3,07495	Значимо
Познание нового	0,1458	0,1316	0,245467	Незначимо
Активная жизнь	0,0833	0,4868	−7,09485	Значимо
Возможность творчества	0,0417	0,0329	0,272634	Незначимо

Итак, можно констатировать определенные изменения ценностного сознания с возрастом.

8. Для сравнения двух ранжированных рядов в целом рассчитаем коэффициент ранговой корреляции Спирмена и запишем результаты в табличной форме.

Ценности	Моло- дежь	Взрос- лые	Моло- дежь	Взрос- лые	$d_i$	$d_i^2$
	Частоты		Ранг			
Счастливая семейная жизнь	35	123	1,5	2	−0,5	0,25
Здоровье	35	113	1,5	3	−1,5	2,25
Материально обеспечен- ная жизнь	31	142	3	1	2	4
Интересная работа	27	56	4	5	−1	1
Любовь	25	11	5	10	−5	25
Наличие хороших и вер- ных друзей	21	47	6	7	−1	1
Карьера	11	6	7	11	−4	16
Удовольствия	10	16	8	9	−1	1
Общественное признание	7	52	9,5	6	3,5	12,25
Познание нового	7	20	9,5	8	1,5	2,25
Активная жизнь	4	74	11	4	7	49
Возможность творчества	2	5	12	12	0	0
Итого	215	665	—	—	—	114

Если две ценности имеют одинаковые частоты, то каждой из них приписываем ранг, равный полусумме тех соседних рангов, которые бы получили они, оставаясь на прежних местах ранжированного ряда, но имея при этом чуть разнящиеся частоты. Так, две первые ценности в молодежном ряде, имея одинаковые частоты (35), делят между собой первый и второй ранги. Следовательно, каждая из них получает в этом ряду ранг, равный 1,5. Следующая за ними ценность получает третий ранг. Если более двух ценностей имеют одинаковые частоты, то точно так же каждой из них приписываем ранг, равный среднему арифметическому тех рангов, которые они делят между собой.

Формула коэффициента ранговой корреляции Спирмена:

$$r_s = 1 - 6 \frac{\sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

где  $d_i$  — разница между парой рангов,  $n$  — количество сравниваемых пар рангов.

Значение коэффициента ранговой корреляции Спирмена находится между минус единицей и единицей:  $-1 \leq r_s \leq 1$ .

Подсчитаем коэффициент Спирмена:

$$r_{\text{эмп}} = 1 - 6 \frac{114}{12(12^2 - 1)} = 0,601.$$

Незначительное число сравниваемых пар даже при относительно высоком по модулю коэффициенте Спирмена не дает оснований утверждать наличие корреляционной зависимости ранжированных рядов в генеральной совокупности. Необходимо определить значимость полученного коэффициента ранговой корреляции. Возможно, его отличие от нуля обусловлено погрешностями выборочного метода.

Проверим гипотезу о несущественности корреляционной связи между ранжированными рядами в генеральной совокупности,

то есть о случайном отклонении коэффициента ранговой корреляции от нуля в выборочной совокупности. Нулевая гипотеза  $H_0: r_s = 0$ . Альтернативная гипотеза  $H_1: r_s \neq 0$ . Критическая область — двусторонняя. Для надежности  $\gamma = 0,99$  уровень значимости (вероятность ошибки)  $\alpha = 0,01$ .

Критическое значение коэффициента Спирмена найдем по табл. П. 2 (см. приложение):

$$r_{\text{скр}} = 0,727 \text{ при } \alpha = 0,01 \text{ и } n = 12.$$

Если  $|r_{\text{сэмпл}}| > r_{\text{скр}}$ , то значение коэффициента корреляции считается статистически существенным. Это значит, что между рядами существует значимая ранговая корреляционная связь. При этом если  $r_{\text{сэмпл}} > 0$ , то связь прямая (положительная), т. е. высоким рангам показателей одного ряда будут соответствовать также высокие ранги аналогичных показателей другого ряда, а низким — низкие. Если  $r_{\text{сэмпл}} < 0$ , то связь обратная (отрицательная), т. е. показателям с высокими рангами одного ряда будут соответствовать низкие ранги этих же показателей в другом ряде и наоборот.

Если  $|r_{\text{сэмпл}}| < r_{\text{скр}}$ , то принимают гипотезу о несущественности ранговой корреляционной связи. Чем ближе к нулю абсолютная величина  $r_{\text{сэмпл}}$ , тем меньше зависимость между рядами. При отсутствии корреляции все ранги показателей в разных рядах перемешаны, между ними нет никакого соответствия.

Так как в нашем случае значение  $|r_{\text{сэмпл}}| < r_{\text{скр}}$ , то ранговой корреляции между рядами нет, т. е. взаимосвязь ранжированных рядов не значима.

Значимость коэффициента ранговой корреляции Спирмена можно также проверить по значению величины  $T_{\text{кр}} = t_{\alpha; n-2} \sqrt{(1 - r_{\text{сэмпл}}^2) / (n - 2)}$ , где  $t_{\alpha; n-2}$  — критическая точка распределения Стьюдента для двусторонней критической области при уровне значимости  $\alpha$  и числе степеней свободы  $k = n - 2$ . Если  $|r_{\text{сэмпл}}| < T_{\text{кр}}$ , то связь между признаками незначима (при уровне

значимости  $\alpha$ ). Если  $|r_{\text{эмп}}| > T_{\text{кр}}$ , то между признаками существует значимая ранговая корреляционная связь.

В нашем случае  $\alpha = 0,01$ ,  $k = n - 2 = 12 - 2 = 10$ ,  $t_{\alpha; n-2} = t_{0,01; 10} = 3,17$ .

$$T_{\text{кр}} = 3,17 \sqrt{(1 - 0,601^2) / 10} = 0,801 > 0,601 = r_s, \text{ т. е. } |r_s| < T_{\text{кр}}.$$

Следовательно, с надежностью 0,99 (с вероятностью ошибки 0,01) ранговая связь между признаками незначима.

Статистический вывод: с вероятностью ошибки менее 0,01 можно утверждать, что предпочитаемые ценности в разных возрастных группах отличаются, но ранжированные ряды не противоположны ( $r_{\text{эмп}} > 0$ ). Следовательно, имеет смысл сравнивать степень предпочтения отдельных ценностей, применяя, помимо критерия Стьюдента, другие статистические показатели.

Упорядочим теперь и вторую группу респондентов («взрослые») по убыванию частоты выбора, т. е. по снижению статуса ценности. Суббыванием значимости ценности ее ранговый показатель увеличивается.

#### Сравнение ранжированных рядов

Молодежь		Взрослые	
Ценности	Ранг	Ценности	Ранг
Счастливая семейная жизнь	1–2	Материально обеспеченная жизнь	1
Здоровье	1–2	Счастливая семейная жизнь	2
Материально обеспеченная жизнь	3	Здоровье	3
Интересная работа	4	Активная жизнь	4
Любовь	5	Интересная работа	5
Наличие хороших и верных друзей	6	Общественное признание	6
Карьера	7	Наличие хороших и верных друзей	7
Удовольствия	8	Познание нового	8
Общественное признание	9–10	Удовольствия	9
Познание нового	9–10	Любовь	10
Активная жизнь	11	Карьера	11
Возможность творчества	12	Возможность творчества	12

При сравнении полученных рядов можно заметить, что со вступлением во взрослую жизнь значимость таких ценностей, как материально обеспеченная жизнь, активная жизнь, общественное признание повышается, а статус таких ценностей, как любовь, карьера, удовольствия снижется. В то же время стоит отметить, что материально обеспеченная жизнь, счастливая семейная жизнь и здоровье занимают три первых места в ряду ценностей в обеих группах. Близкие по рангу места занимают интересная работа и наличие хороших и верных друзей. Несколько удивляет более низкий статус ценности познание нового в группе молодежи. Это может быть связано с изменением ценностной парадигмы поколений: с большим рационализмом современного молодого и большим романтизмом старшего поколения. Но возможно, такая разница является следствием выборочного метода. Следует проверить, не находится ли эта разность в пределах ошибки выборки.



# ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

---

## *Задание № 1*

**Тема:** анализ одномерных распределений.

**Цель:** отработать навыки математической обработки и анализа массива социологических данных.

**Задание.** Рассмотрите представленную таблицу одномерного распределения. Рассчитайте статистические показатели, пригодные для данных, замеренных на этом уровне.

Порядок выполнения:

- 1) определить вид переменной и шкалу измерений;
- 2) определить показатели, пригодные для данной шкалы: для каждого признака найти процент от числа опрошенных, от числа ответивших, от числа ответов, дисперсию доли, предельную ошибку выборки с надежностью 0,95;
- 3) осуществить расчет выбранных показателей. Свести результаты в таблицу, упорядоченную по частоте ответов;
- 4) провести социологический анализ полученных результатов.

**Предпочитаемые способы проведения досуга по опросу молодежи****Вариант 1**

Способы проведения досуга	Абсолютные значения
Общение с друзьями	28
Посещение кино	16
Дискотека	42
Общение в интернете	31
Посещение театров	14
Просмотр телесериалов	27
Занятия спортом	14
Турпоходы	25
Пикники на природе	53
Посещение филармонии	7
Посещение концертов	26
Занятия танцами	11
Занятия рисованием и живописью	3
Чтение книг	5
Посещение художественных выставок	4
Коллекционирование	3
Участие в коллективах самодеятельности	1
Итого	310

Количество опрошенных — 300 человек.

Не ответили на данный вопрос 25 человек.

**Вариант 2**

Способы проведения досуга	Абсолютные значения
Общение с друзьями	25
Посещение кино	20
Дискотека	34
Общение в интернете	57
Посещение театров	28
Просмотр телесериалов	32
Занятия спортом	16
Турпоходы	12
Пикники на природе	33
Посещение филармонии	17

Способы проведения досуга	Абсолютные значения
Посещение концертов	23
Занятия танцами	19
Занятия рисованием и живописью	13
Чтение книг	15
Посещение художественных выставок	9
Коллекционирование	7
Участие в коллективах самодеятельности	5
Итого	365

Количество опрошенных — 320 человек.

Не ответили на данный вопрос 27 человек.

### Вариант 3

Способы проведения досуга	Абсолютные значения
Общение с друзьями	19
Посещение кино	26
Дискотека	72
Общение в интернете	37
Посещение театров	24
Просмотр телесериалов	47
Занятия спортом	31
Турпоходы	25
Пикники на природе	42
Посещение филармонии	17
Посещение концертов	28
Занятия танцами	14
Занятия рисованием и живописью	21
Чтение книг	17
Посещение художественных выставок	25
Коллекционирование	7
Участие в коллективах самодеятельности	10
Итого	462

Количество опрошенных — 400 человек.

Не ответили на данный вопрос 26 человек.

#### Вариант 4

Способы проведения досуга	Абсолютные значения
Общение с друзьями	46
Посещение кино	28
Дискотека	37
Общение в интернете	61
Посещение театров	14
Просмотр телесериалов	30
Занятия спортом	18
Турпоходы	25
Пикники на природе	43
Посещение филармонии	7
Посещение концертов	16
Занятия танцами	9
Занятия рисованием и живописью	14
Чтение книг	25
Посещение художественных выставок	8
Коллекционирование	3
Участие в коллективах самодеятельности	12
Итого	396

Количество опрошенных — 350 человек.

Не ответили на данный вопрос 32 человека.

#### Вариант 5

Способы проведения досуга	Абсолютные значения
Общение с друзьями	20
Посещение кино	36
Дискотека	52
Общение в интернете	51
Посещение театров	24
Просмотр телесериалов	47
Занятия спортом	21
Турпоходы	32
Пикники на природе	73
Посещение филармонии	14
Посещение концертов	27
Занятия танцами	16

Способы проведения досуга	Абсолютные значения
Занятия рисованием и живописью	10
Чтение книг	35
Посещение художественных выставок	11
Коллекционирование	13
Участие в коллективах самодеятельности	8
Итого	490

Количество опрошенных — 450 человек.

Не ответили на данный вопрос 31 человек.

#### Вариант 6

Способы проведения досуга	Абсолютные значения
Общение с друзьями	22
Посещение кино	16
Дискотека	24
Общение в интернете	33
Посещение театров	14
Просмотр телесериалов	27
Занятия спортом	14
Турпоходы	15
Пикники на природе	13
Посещение филармонии	5
Посещение концертов	10
Занятия танцами	11
Занятия рисованием и живописью	3
Чтение книг	5
Посещение художественных выставок	8
Коллекционирование	3
Участие в коллективах самодеятельности	2
Итого	225

Количество опрошенных — 206 человек.

Не ответили на данный вопрос 11 человек.

Вариант 7

Способы проведения досуга	Абсолютные значения
Общение с друзьями	26
Посещение кино	20
Дискотека	14
Общение в интернете	21
Посещение театров	28
Просмотр телесериалов	37
Занятия спортом	10
Турпоходы	12
Пикники на природе	13
Посещение филармонии	17
Посещение концертов	23
Занятия танцами	9
Занятия рисованием и живописью	13
Чтение книг	4
Посещение художественных выставок	9
Коллекционирование	7
Участие в коллективах самодеятельности	5
Итого	268

Количество опрошенных — 280 человек.

Не ответили на данный вопрос 15 человек.

Вариант 7

Способы проведения досуга	Абсолютные значения
Общение с друзьями	39
Посещение кино	23
Дискотека	50
Общение в интернете	17
Посещение театров	22
Просмотр телесериалов	37
Занятия спортом	24
Турпоходы	15
Пикники на природе	22
Посещение филармонии	17
Посещение концертов	16
Занятия танцами	14

Способы проведения досуга	Абсолютные значения
Занятия рисованием и живописью	21
Чтение книг	14
Посещение художественных выставок	25
Коллекционирование	11
Участие в коллективах самодеятельности	10
Итого	377

Количество опрошенных — 390 человек.

Не ответили на данный вопрос 26 человек.

#### Вариант 9

Способы проведения досуга	Абсолютные значения
Общение с друзьями	21
Посещение кино	38
Дискотека	27
Общение в интернете	44
Посещение театров	14
Просмотр телесериалов	23
Занятия спортом	18
Турпоходы	15
Пикники на природе	33
Посещение филармонии	12
Посещение концертов	16
Занятия танцами	9
Занятия рисованием и живописью	14
Чтение книг	25
Посещение художественных выставок	8
Коллекционирование	3
Участие в коллективах самодеятельности	12
Итого	332

Количество опрошенных — 295 человек.

Не ответили на данный вопрос 12 человек.

## Вариант 10

Способы проведения досуга	Абсолютные значения
Общение с друзьями	29
Посещение кино	30
Дискотека	54
Общение в интернете	31
Посещение театров	24
Просмотр телесериалов	43
Занятия спортом	21
Турпоходы	12
Пикники на природе	33
Посещение филармонии	14
Посещение концертов	37
Занятия танцами	11
Занятия рисованием и живописью	9
Чтение книг	15
Посещение художественных выставок	19
Коллекционирование	13
Участие в коллективах самодеятельности	3
Итого	398

Количество опрошенных — 420 человек.

Не ответили на данный вопрос 11 человек.



## Задание № 2

**Тема:** измерение, группировка, отображение и анализ первичных социометрических данных.

**Цель:** отработать навыки математической обработки и анализа массива социологических данных.

**Задание.** Определите тип переменной и упорядочите данную совокупность показателей социологического исследования. Постройте частотный вариационный ряд, полигон и эмпирическую функцию распределения. Найдите среднее значение переменной, моду и медиану.

Ниже приведены данные исследования жилищных условий группы работающей молодежи. На вопрос о количестве комнат в квартире проживания были получены следующие ответы.

## Вариант 1

3 1 4 2 3 2 2 2 4 3 2 4 2 3 4 3 2 3 1 3 4 1 2 3 4  
 3 2 3 4 2 3 2 4 1 3 2 2 3 2 4 2 2 3 1 2 3 3 2 3  
 1 4 3 2 2 4 1 2 2 4 3 2 1 2 3 4 3 4 2 2 2 3 2 2  
 4 1 3 4 3 4 3 4 2 1 3 4 1 2 1 3 2 3 3 4 2 3 2 1

## Вариант 2

2 1 1 2 4 2 5 2 4 3 2 4 2 3 4 3 2 3 1 3 4 1 2 3 4  
 3 2 4 4 2 3 2 4 1 3 2 2 3 2 5 2 4 3 1 2 3 3 2 3 3  
 2 4 3 2 2 4 1 2 1 4 3 2 1 2 3 4 1 4 2 1 2 3 1 3  
 5 1 3 4 1 3 3 4 2 1 3 4 1 1 1 3 4 3 3 4 2 3 2 2

## Вариант 3

4 2 4 2 3 2 3 2 4 3 2 4 2 3 4 3 2 3 1 3 4 1 2 3  
 3 2 3 4 1 3 2 4 1 3 1 2 3 2 3 2 2 3 1 3 3 3 4 3  
 2 4 2 2 2 4 1 2 4 4 3 2 1 2 3 4 1 4 2 2 2 3 1 2  
 4 1 3 4 3 4 3 4 2 1 3 4 1 1 1 3 2 2 3 4 2 4 2 1

### Вариант 4

3 1 4 2 2 2 1 2 4 3 2 4 2 3 4 3 2 3 1 3 4 1 2 3 4  
 1 2 4 4 2 3 2 4 1 3 3 2 3 1 4 2 2 3 1 1 3 3 2 3 3  
 1 4 2 2 1 4 3 2 2 4 3 2 1 2 2 4 3 4 2 2 2 2 1 2 1  
 2 1 3 4 3 4 3 4 2 1 3 4 1 1 1 3 2 3 3 4 2 3 2 1 2

### Вариант 5

3 3 1 2 3 2 1 2 4 3 2 4 2 3 4 2 2 3 1 3 4 1 2 3 4  
 4 2 3 4 2 3 2 4 1 3 5 2 3 2 4 2 2 3 1 2 3 3 2 3 2  
 1 2 3 2 3 4 4 2 2 4 3 2 1 2 5 4 3 4 2 2 2 3 5 2 2  
 1 1 3 4 3 4 3 4 2 1 3 4 1 1 1 3 2 1 3 4 2 3 3 1

### Вариант 6

3 2 4 2 3 2 1 2 4 3 2 4 2 3 4 3 2 3 1 3 4 2 2 3 4  
 3 2 3 4 2 3 2 4 1 3 2 2 3 2 4 2 2 3 2 2 3 3 2 4  
 1 4 3 2 2 4 1 2 2 4 3 2 1 2 2 4 3 4 2 2 2 3 1 2  
 4 1 2 4 3 4 2 4 2 2 3 4 1 1 1 3 2 3 3 3 2 3 2 1

### Вариант 7

2 1 3 2 4 2 5 2 4 3 2 4 2 3 4 3 2 3 1 3 4 5 2 3 4  
 3 2 4 4 2 3 2 4 5 3 2 2 3 2 5 2 4 3 1 2 3 3 2 3 3  
 2 4 3 2 2 5 5 2 1 4 3 2 3 2 3 4 1 4 2 3 2 3 1 3  
 5 1 3 4 3 3 3 4 2 1 3 5 1 2 1 3 4 3 3 4 2 3 2 2

### Вариант 8

4 2 4 2 3 2 3 2 4 3 2 4 2 3 4 3 2 3 1 3 4 2 2 3  
 3 2 3 4 2 3 2 4 2 3 1 2 3 2 3 2 2 3 2 3 3 3 4 3  
 2 4 2 2 2 3 1 2 2 4 3 2 2 2 3 4 1 2 2 2 2 3 1 2  
 4 2 3 4 3 4 3 4 2 2 3 4 1 2 2 3 2 2 3 3 2 4 2 1

Вариант 9

3 1 4 2 2 2 3 2 4 3 2 4 2 3 4 3 2 3 1 3 4 1 2 3 4  
3 2 4 4 2 3 2 4 2 3 3 2 3 1 3 2 2 3 2 1 3 3 2 3 3  
1 3 2 2 2 4 3 2 2 4 3 2 2 2 2 4 3 3 2 2 2 2 2 2 1  
2 1 3 4 3 4 3 4 2 2 3 4 1 2 2 3 2 3 3 4 2 3 2 3 2

Вариант 10

3 3 1 2 3 2 2 2 4 3 2 4 2 3 4 2 2 3 2 3 4 1 2 3 5  
4 2 3 4 2 3 2 5 1 3 5 2 3 2 4 2 2 3 2 2 3 3 2 3 2  
2 2 3 2 3 2 4 2 2 3 3 2 2 2 3 4 3 4 2 2 2 3 5 2 2  
1 2 3 4 3 4 3 4 2 2 3 4 2 2 1 3 2 2 3 4 2 3 3 1

### Задание № 3

**Тема:** измерение, группировка, отображение и анализ первичных социометрических данных.

**Цель:** отработать навыки группировки, математической обработки и графического отображения первичных социометрических данных.

**Задание.** Сгруппируйте данные социологического исследования, разбив их совокупность на пять интервалов. Постройте интервальный вариационный ряд, найдите размах вариации, среднее значение переменной, моду и медиану. Постройте гистограмму и кумуляту, полигон и эмпирическую функцию распределения.

Ниже приведены данные опроса группы студентов о среднем времени, которое они затрачивают, чтобы добраться до места учебы. Были получены следующие ответы (в мин).

#### Вариант 1

37 24 20 16 18 32 15 14 22 24 14 19 16 18 16 14 15 27 26 16  
13 29 27 19 18 14 16 17 26 18 16 52 12 13 18 24 12 18 28 21  
38 45 36 12 26 23 25 44 48 21 17 12 15 29 39 20 25 57 21 30  
35 27 21 10 18 47 13 38 26 26 11 26 32 51 23 19 16 23 14  
12 23 11 15 13 22 27 15 42 27 14 13 14 10 28 53 13 20 17

#### Вариант 2

27 31 20 16 19 32 11 14 20 24 15 19 17 18 26 14 13 27 28 16  
11 29 23 19 12 14 14 17 27 18 19 52 18 13 13 24 22 18 20 21  
37 43 36 16 26 28 25 34 48 41 17 22 15 39 39 25 25 52 21  
31 27 24 10 18 45 13 38 29 26 11 22 32 51 27 19 16 23 14  
12 23 11 15 33 22 27 15 32 27 14 13 14 10 21 53 13 20 17

### Вариант 3

17 24 24 16 17 32 14 14 21 24 27 19 15 18 16 14 13 27 26 16  
 13 29 27 19 18 14 16 17 26 18 16 52 12 13 14 24 18 18 28  
 28 45 32 12 25 23 23 44 43 21 14 12 12 29 39 20 25 57 21  
 35 27 21 10 18 47 13 38 26 26 11 26 32 51 33 19 26 23 14  
 32 23 16 15 17 22 29 15 12 27 18 13 16 10 28 53 13 20 17

### Вариант 4

37 28 20 19 18 22 15 16 22 14 12 19 16 18 16 14 15 27 26 16  
 13 29 27 13 18 24 16 17 26 28 16 52 14 13 19 24 12 28 21 21  
 38 35 36 11 26 13 25 24 48 21 17 12 15 29 39 26 25 57 21 30  
 35 26 21 17 18 42 13 48 26 24 11 26 31 51 13 19 16 23 34 42  
 12 13 11 14 13 21 27 15 42 27 16 13 44 10 28 53 13 20 17

### Вариант 5

27 24 30 16 28 32 15 14 22 24 14 19 16 18 16 14 15 27 26 16  
 42 29 27 19 18 14 14 17 23 18 16 52 12 53 18 24 12 18 28 21  
 34 45 36 12 16 29 25 44 48 31 17 32 15 29 39 20 25 57 21 30  
 35 25 21 10 18 47 13 38 26 26 41 26 32 51 23 19 16 23 14 26  
 12 33 11 15 13 22 27 18 42 27 14 13 44 10 28 53 13 20 17 13

### Вариант 6

37 24 20 16 18 32 15 21 22 24 14 19 16 18 24 14 15 27 26 16  
 20 29 27 19 18 23 16 17 26 18 16 52 12 13 18 24 22 18 28 21  
 38 45 36 12 26 23 25 44 48 21 17 12 26 29 39 20 25 57 21 30  
 35 27 21 28 18 47 13 38 26 26 11 26 32 51 23 19 16 23 14  
 12 23 25 15 13 22 27 27 42 27 29 13 14 10 28 53 13 20 17

### Вариант 7

27 31 20 16 19 32 31 14 20 24 15 49 17 38 26 14 33 27 28 16  
 11 29 23 19 32 14 34 47 27 18 19 52 38 13 33 24 22 18 20 21  
 37 43 36 16 26 28 25 34 48 41 37 22 15 39 39 25 25 52 21  
 31 27 24 10 18 45 43 38 29 26 41 22 32 51 27 19 36 23 44  
 42 23 11 15 33 22 27 45 32 27 34 13 44 10 21 53 43 20 47

### Вариант 8

17	24	24	36	47	32	54	34	21	24	27	59	45	38	16	34	43	27	26	36
13	29	27	19	18	14	56	17	26	18	16	52	52	33	54	24	48	18	28	
28	45	32	12	25	23	23	44	43	21	14	32	12	29	39	20	25	27	21	
35	27	21	10	48	47	53	38	26	26	31	26	32	51	33	49	26	23	44	
32	23	46	15	57	22	29	55	42	27	18	33	46	30	28	23	53	20	57	

### Вариант 9

37	28	20	19	18	22	35	16	42	14	32	19	36	18	16	34	15	27	26	16
13	29	27	13	18	24	16	17	26	28	16	52	14	13	19	24	12	48	21	41
38	35	36	11	26	13	45	24	48	21	37	12	15	29	39	26	25	57	41	30
35	26	41	17	18	42	33	48	26	24	31	46	31	51	13	19	36	43	34	42
12	13	11	34	13	41	27	15	42	47	16	33	44	10	28	53	13	20	17	

### Вариант 10

27	44	30	16	28	32	15	14	42	54	14	19	16	18	36	14	15	47	26	16
42	29	57	19	18	14	14	17	43	18	16	52	12	53	18	54	12	18	48	41
34	45	36	12	16	59	25	44	48	31	17	32	15	29	39	40	25	57	21	30
35	25	41	10	18	47	33	38	46	26	41	46	32	51	43	19	16	43	14	46
12	33	11	15	13	42	27	38	42	47	14	13	44	10	58	53	13	20	17	13

## Задание № 4

**Тема:** выборочный метод в социологии

**Цель:** отработка методов математического анализа количественных данных социологического опроса.

**Задание.** Проведите анализ представленных выборочных данных по заданному алгоритму с целью определения показателей генеральной совокупности.

**Порядок выполнения:**

- 1) определить среднее значение показателя в генеральной совокупности, гарантируя результат с заданной надежностью;
- 2) определить долю элементов генеральной совокупности, обладающих заданным значением показателя, гарантируя результат с заданной надежностью;
- 3) найти необходимую численность выборки при определении среднего значения показателя, чтобы с заданной надежностью предельная ошибка выборки не превышала заданного значения;
- 4) найти необходимую численность выборки при определении доли элементов генеральной совокупности, обладающих заданным значением показателя, чтобы с заданной надежностью предельная ошибка выборки не превышала заданного значения.

## Вариант 1

В порядке случайной бесповторной выборки было опрошено 330 из 3300 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 2000	2000–3000	3000–4000	4000 и более
Число студентов	18	145	121	46

**Определить:**

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 95,4 %;

— долю подрабатывающих студентов, имеющих месячный доход 3000 руб. и выше, гарантируя результат с вероятностью 0,950.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 100 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 3000 руб. и выше, чтобы с надежностью 99 % предельная ошибка не превышала 5 %.

### Вариант 2

В порядке случайной бесповторной выборки было опрошено 250 из 2000 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 3000	3000–3500	3500–4000	4000 и более
Число студентов	21	105	91	33

*Определить:*

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 99,7 %;

— долю подрабатывающих студентов, имеющих месячный доход 3500 руб. и выше, гарантируя результат с вероятностью 0,950.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,99 предельная ошибка выборки не превышала 200 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 3500 руб. и выше, чтобы с надежностью 95,4 % предельная ошибка не превышала 3 %.



## Вариант 3

На предприятии в порядке случайной бесповторной выборки было опрошено 150 молодых сотрудников из 1500. Получены следующие данные об их доходе за октябрь:

Месячный доход, руб.	Менее 2000	2000–2500	2500–3000	3000 и выше
Число сотрудников	22	80	30	18

*Определить:*

- среднемесячный размер дохода у молодых сотрудников данного предприятия, гарантируя результат с вероятностью 0,950;
- долю сотрудников предприятия, имеющих месячный доход 2500 руб. и выше, гарантируя результат с надежностью 95,4 %.

*Найти:*

- необходимую численность выборки при определении среднего месячного дохода сотрудников предприятия, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 150 руб.;
- необходимую численность выборки при определении доли сотрудников с размером месячного дохода 2500 руб. и выше, чтобы с надежностью 99 % предельная ошибка не превышала 4 %.

## Вариант 4

В порядке случайной бесповторной выборки было опрошено 350 из 2000 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 2000	2000–4000	4000–6000	6000 и более
Число студентов	28	139	147	36

*Определить:*

- среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 95 %;

— долю подрабатывающих студентов, имеющих месячный доход 4000 руб. и выше, гарантируя результат с вероятностью 0,954.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 300 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 4000 руб. и выше, чтобы с надежностью 99 % предельная ошибка не превышала 2 %.

#### Вариант 5

В порядке случайной бесповторной выборки было опрошено 250 из 1500 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 3500	3500—4000	4000—4500	4500 и более
Число студентов	21	115	90	24

*Определить:*

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 99 %;

— долю подрабатывающих студентов, имеющих месячный доход 3500 руб. и выше, гарантируя результат с вероятностью 0,954.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 250 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 3500 руб. и выше, чтобы с надежностью 95 % предельная ошибка не превышала 6 %.

## Вариант 6

В порядке случайной бесповторной выборки было опрошено 500 из 5000 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 4000	4000–4500	4500–5000	5000 и более
Число студентов	180	150	120	50

*Определить:*

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 95,4 %;

— долю подрабатывающих студентов, имеющих месячный доход 4000 руб. и выше, гарантируя результат с вероятностью 0,950.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 200 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 4000 руб. и выше, чтобы с надежностью 99 % предельная ошибка не превышала 5 %.

## Вариант 7

В порядке случайной бесповторной выборки было опрошено 250 из 2000 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 3500	3500–4000	4000–4500	4500 и более
Число студентов	21	85	90	54

*Определить:*

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 99,7 %;

— долю подрабатывающих студентов, имеющих месячный доход 3500 руб. и выше, гарантируя результат с вероятностью 0,950.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,99 предельная ошибка выборки не превышала 100 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 3500 руб. и выше, чтобы с надежностью 95,4 % предельная ошибка не превышала 3 %.

### Вариант 8

На предприятии в порядке случайной бесповторной выборки было опрошено 350 молодых сотрудников из 3500. Получены следующие данные об их доходе за октябрь:

Месячный доход, руб.	Менее 2000	2000—3000	3000—4000	4000 и выше
Число сотрудников	72	130	110	38

*Определить:*

— среднемесячный размер дохода у молодых сотрудников данного предприятия, гарантируя результат с вероятностью 0,950;

— долю сотрудников предприятия, имеющих месячный доход 3000 руб. и выше, гарантируя результат с надежностью 95,4 %.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода сотрудников предприятия, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 150 руб.;

— необходимую численность выборки при определении доли сотрудников с размером месячного дохода 3000 руб. и выше, чтобы с надежностью 99 % предельная ошибка не превышала 4 %.

## Вариант 9

В порядке случайной бесповторной выборки было опрошено 400 из 2000 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 2000	2000–4000	4000–6000	6000 и более
Число студентов	48	159	147	46

*Определить:*

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 95 %;

— долю подрабатывающих студентов, имеющих месячный доход 4000 руб. и выше, гарантируя результат с вероятностью 0,954.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 300 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 4000 руб. и выше, чтобы с надежностью 99 % предельная ошибка не превышала 2 %.

## Вариант 10

В порядке случайной бесповторной выборки было опрошено 300 из 1500 подрабатывающих студентов. Получены следующие данные об их доходе за текущий месяц:

Месячный доход, руб.	Менее 3500	3500–4000	4000–4500	4500 и более
Число студентов	51	135	80	34

*Определить:*

— среднемесячный размер дохода у подрабатывающих студентов, гарантируя результат с надежностью 99 %;

— долю подрабатывающих студентов, имеющих месячный доход 3500 руб. и выше, гарантируя результат с вероятностью 0,954.

*Найти:*

— необходимую численность выборки при определении среднего месячного дохода подрабатывающих студентов, чтобы с вероятностью 0,997 предельная ошибка выборки не превышала 150 руб.;

— необходимую численность выборки при определении доли подрабатывающих студентов с размером месячного дохода 3500 руб. и выше, чтобы с надежностью 95 % предельная ошибка не превышала 1 %.

## Задание № 5

**Тема:** анализ модели поведения социальной группы.

**Цель:** отработка методов математического анализа модели группового поведения.

**Задание.** Представлены данные о частоте того или иного поведения представителей определенных социальных групп. Проведите сравнительный анализ поведения различных групп исследуемой совокупности и проинтерпретируйте результаты.

Порядок выполнения:

1) определить условные индексы для каждой группы и в целом по массиву данных (по формуле средней арифметической взвешенной);

2) рассчитать предельную ошибку выборки для определения значимости различий между группами с надежностью 0,95;

3) построить сравнительную диаграмму, найти моду признака для каждой группы и по массиву в целом;

4) провести социологический анализ полученных результатов.

Ниже представлены данные о предпочтительном способе проведения досуга студентами одного из вузов в зависимости от пола.

Частота выбора одного из видов проведения досуга  
как предпочтительного  
(в абсолютных числах)

## Вариант 1

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	197	92	105
Общаюсь с родными	86	25	61
Общаюсь с друзьями	79	32	47
Читаю книги, журналы, газеты	39	20	19
Слушаю музыку	132	75	57

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Хожу на дискотеки	137	48	89
Играю в компьютерные игры	155	113	42
Затрудняюсь ответить	24	15	9
Итого	849	420	429

### Вариант 2

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	200	95	105
Общаюсь с родными	76	25	51
Общаюсь с друзьями	89	37	52
Читаю книги, журналы, газеты	39	20	19
Слушаю музыку	121	65	56
Хожу на дискотеки	148	58	90
Играю в компьютерные игры	135	103	32
Затрудняюсь ответить	44	25	19
Итого	852	428	424

### Вариант 3

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	256	102	154
Общаюсь с родными	164	75	89
Общаюсь с друзьями	100	43	57
Читаю книги, журналы, газеты	99	30	69
Слушаю музыку	173	96	77
Хожу на дискотеки	127	38	89
Играю в компьютерные игры	190	128	62
Затрудняюсь ответить	17	11	6
Итого	1126	523	603



## Вариант 4

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	179	84	95
Общаюсь с родными	93	35	58
Общаюсь с друзьями	139	67	72
Читаю книги, журналы, газеты	99	40	59
Слушаю музыку	120	67	53
Хожу на дискотеки	136	51	85
Играю в компьютерные игры	115	93	22
Затрудняюсь ответить	24	15	9
Итого	905	452	453

## Вариант 5

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	236	111	125
Общаюсь с родными	112	45	67
Общаюсь с друзьями	146	52	94
Читаю книги, журналы, газеты	91	32	59
Слушаю музыку	152	85	67
Хожу на дискотеки	86	47	39
Играю в компьютерные игры	134	92	42
Затрудняюсь ответить	36	17	19
Итого	993	481	512

## Вариант 6

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	169	73	96
Общаюсь с родными	51	20	31
Общаюсь с друзьями	103	48	55
Читаю книги, журналы, газеты	20	11	9

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Слушаю музыку	157	83	74
Хожу на дискотеки	141	52	89
Играю в компьютерные игры	109	67	42
Затрудняюсь ответить	13	5	8
Итого	763	359	404

### Вариант 7

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	131	42	89
Общаюсь с родными	78	31	47
Общаюсь с друзьями	135	67	68
Читаю книги, журналы, газеты	41	12	29
Слушаю музыку	151	75	76
Хожу на дискотеки	113	58	55
Играю в компьютерные игры	118	96	22
Затрудняюсь ответить	15	5	10
Итого	782	386	396

### Вариант 8

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	236	132	104
Общаюсь с родными	135	56	79
Общаюсь с друзьями	160	73	87
Читаю книги, журналы, газеты	77	54	23
Слушаю музыку	198	102	96
Хожу на дискотеки	156	67	89
Играю в компьютерные игры	131	99	32
Затрудняюсь ответить	16	10	6
Итого	1109	593	516

## Вариант 9

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	251	134	117
Общаюсь с родными	180	95	85
Общаюсь с друзьями	183	87	96
Читаю книги, журналы, газеты	85	48	37
Слушаю музыку	160	77	83
Хожу на дискотеки	159	61	98
Играю в компьютерные игры	170	108	62
Затрудняюсь ответить	21	12	9
Итого	1209	622	587

## Вариант 10

Способ проведения досуга	В целом по массиву	Пол	
		Юноши	Девушки
Смотрю телевизор	261	93	168
Общаюсь с родными	176	82	94
Общаюсь с друзьями	183	74	109
Читаю книги, журналы, газеты	52	13	39
Слушаю музыку	212	115	97
Хожу на дискотеки	185	97	88
Играю в компьютерные игры	155	123	32
Затрудняюсь ответить	19	7	12
Итого	1243	604	639

## Задание № 6

**Тема:** сравнительный анализ распределения общественного мнения в разных социальных группах.

**Цель:** отработка математических методов анализа общественного мнения.

**Задание.** По представленным данным ранжированной оценки системы высказываний проанализируйте распределение общественного мнения в разных социальных группах.

Порядок выполнения:

- 1) рассчитать условные индексы для каждой группы (по формуле средней арифметической взвешенной);
- 2) рассчитать ошибку выборки, оценить значимость различий между значениями индекса в разных группах с надежностью 0,954;
- 3) представить результаты в графической форме;
- 4) сделать социологический вывод относительно полученных результатов.

Был проведен опрос группы молодежи по следующей анкете:

«Вам предложены пары противоположных высказываний. Выберите то, с которым вы согласны, и отметьте, насколько. Обведите кружком по одной цифре в каждой строке».

Счастливым можно назвать только богатого человека	3	2	1	0	1	2	3	Для счастья богатство не важно; и в бедности человек может быть счастливым
Я не представляю свою жизнь без друзей	3	2	1	0	1	2	3	Друзья мне не очень нужны; я смогу прожить и без них
Для меня важно быть уверенным в завтрашнем дне. Жить только настоящим — глупо	3	2	1	0	1	2	3	Нужно жить только настоящим, не задумываясь, что будет завтра

Каждый человек должен жить и трудиться на благо своей Родины	3	2	1	0	1	2	3	Жить надо прежде всего для себя и своих близких
Мне бы не хотелось, чтобы кто-нибудь вмешивался в мою жизнь	3	2	1	0	1	2	3	Я считаю допустимым вмешательство в мою личную жизнь других людей
Достижение определенного статуса — важнейшая цель моей жизни	3	2	1	0	1	2	3	Я не задумываюсь о достижении определенного положения в обществе

Ранжирование велось по нарастанию выраженности признака, т. е. минимальная выраженность оценивалась значением 1, а максимальная — значением 7.

В результате обработки данных получили следующие распределения.

#### Вариант 1

Выраженность признака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в завтрашнем дне	
1	32	54	64	53	36	68
2	26	48	48	55	44	37
3	23	30	30	42	27	45
4	36	43	27	35	51	72
5	41	29	14	23	35	31
6	33	34	28	34	27	5
7	44	27	24	23	15	7
Всего	235	265	235	265	235	265

Выражен- ность при- знака	Каждый человек должен жить и тру- дится на благо своей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивал- ся в мою жизнь		Достижение опре- деленного стату- са — важнейшая цель моей жизни	
1	40	28	66	38	74	32
2	65	44	53	46	30	30
3	32	36	27	55	24	46
4	29	42	48	68	43	50
5	30	39	19	25	30	29
6	12	37	13	21	23	49
7	27	39	9	12	11	29
Всего	235	265	235	265	235	265

### Вариант 2

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уве- ренным в завтраш- нем дне	
1	23	62	35	44	26	73
2	47	27	28	36	48	48
3	18	34	34	23	37	26
4	32	50	52	47	26	37
5	45	23	47	51	53	25
6	54	41	26	32	20	28
7	20	16	17	20	29	16
Всего	239	253	239	253	239	253
Выражен- ность при- знака	Каждый чело- век должен жить и трудится на бла- го своей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивал- ся в мою жизнь		Достижение опре- деленного стату- са — важнейшая цель моей жизни	
	1	52	32	68	42	76
	2	30	42	45	33	50
3	29	36	52	24	42	37
4	42	53	34	56	34	29
5	36	27	27	28	20	46
6	20	32	10	32	11	39
7	30	31	3	38	6	31
Всего	239	253	239	253	239	253

## Вариант 3

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	43	34	54	50	36	55
2	36	58	38	45	44	31
3	18	47	42	32	27	48
4	28	33	27	40	39	75
5	31	19	24	23	25	21
6	44	24	18	37	37	12
7	25	37	22	25	17	10
Всего	225	252	225	252	225	252
Выражен- ность при- знака	Каждый чело- век должен жить и трудится на бла- го своей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивал- ся в мою жизнь		Достижение опре- деленного стату- са — важнейшая цель моей жизни	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	32	25	59	40	64	32
2	53	34	50	36	30	30
3	42	36	32	55	24	26
4	21	47	38	65	40	47
5	34	39	24	25	33	29
6	16	42	15	21	23	49
7	27	29	7	10	11	39
Всего	225	252	225	252	225	252

## Вариант 4

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	47	52	50	43	32	52
2	32	41	41	50	34	37
3	28	36	37	41	37	42
4	46	47	29	38	45	75
5	31	24	18	24	33	21
6	23	32	26	32	29	15
7	14	17	20	21	11	7
Всего	221	249	221	249	221	249

Выражен- ность при- знака	Каждый чело- век должен жить и трудится на бла- го своей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивался в мою жизнь		Достижение опре- делен-ного стату- са — важнейшая цель моей жизни	
1	40	22	56	38	64	32
2	55	34	49	43	32	37
3	36	38	25	55	24	36
4	21	46	50	52	39	50
5	30	33	21	28	28	29
6	14	37	10	21	23	43
7	25	39	10	12	11	22
Всего	221	249	221	249	221	249

Вариант 5

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за- втрашнем дне	
1	43	62	57	58	46	69
2	31	43	48	50	41	46
3	27	37	31	42	25	35
4	39	41	39	31	54	61
5	42	24	20	33	37	35
6	26	32	28	29	25	11
7	37	23	22	19	17	5
Всего	245	262	245	262	245	262
Выражен- ность при- знака	Каждый человек должен жить и тру- дится на благо сво- ей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивал- ся в мою жизнь		Достижение опре- делен-ного стату- са — важнейшая цель моей жизни	
1	45	38	61	35	61	24
2	55	34	53	41	43	35
3	37	36	37	52	34	46
4	39	42	48	62	43	55
5	30	36	19	33	30	32
6	17	37	13	27	24	43
7	22	39	14	12	10	27
Всего	245	262	245	262	245	262



## Вариант 6

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	62	47	74	83	36	68
2	34	63	48	55	54	57
3	53	51	60	46	27	65
4	27	45	27	37	53	72
5	36	33	19	43	45	32
6	48	56	30	44	32	25
7	22	41	24	28	35	17
Всего	282	336	282	336	282	336
Выражен- ность при- знака	Каждый чело- век должен жить и трудится на бла- го своей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивал- ся в мою жизнь		Достижение опре- деленного стату- са — важнейшая цель моей жизни	
1	40	58	76	78	74	52
2	65	64	53	46	37	70
3	52	56	34	55	54	46
4	39	42	48	68	43	51
5	30	39	19	45	30	39
6	29	38	33	31	23	49
7	27	39	19	13	21	29
Всего	282	336	282	336	282	336

## Вариант 7

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	54	75	45	50	26	73
2	67	46	38	46	48	58
3	34	62	44	33	37	36
4	25	37	62	67	56	37
5	51	49	47	51	53	45
6	42	33	26	62	42	58
7	18	27	29	20	29	22
Всего	291	329	291	329	291	329

Выраженность признака	Каждый человек должен жить и трудиться на благо своей Родины		Мне бы не хотелось, чтобы кто-нибудь вмешивался в мою жизнь		Достижение определенного статуса — важнейшая цель моей жизни	
1	52	62	68	58	76	48
2	30	42	45	43	50	53
3	59	46	72	34	42	47
4	42	63	44	66	54	39
5	46	37	27	38	30	56
6	30	48	10	42	23	49
7	32	31	25	48	16	37
Всего	291	329	291	329	291	329

### Вариант 8

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за- втрашнем дне	
1	50	54	34	43	25	32
2	45	38	58	36	34	53
3	32	42	47	18	36	42
4	40	27	33	28	47	21
5	23	24	19	31	39	34
6	37	18	24	44	42	16
7	25	22	37	25	29	27
Всего	252	225	252	225	252	225
Выражен- ность при- знака	Каждый чело- век должен жить и трудится на бла- го своей Родины		Мне бы не хоте- лось, чтобы кто- нибудь вмешивался в мою жизнь		Достижение опре- делен-ного стату- са — важнейшая цель моей жизни	
	1	55	36	32	64	40
2	31	44	30	30	36	50
3	48	27	26	24	55	32
4	75	39	47	40	65	38
5	21	25	29	33	25	24
6	12	37	49	23	21	15
7	10	17	39	11	10	7
Всего	252	225	252	225	252	225

## Вариант 9

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	70	47	50	50	32	42
2	42	41	41	43	54	37
3	23	36	57	41	37	52
4	41	52	29	38	45	75
5	37	24	17	24	33	15
6	13	32	36	32	28	21
7	24	17	20	21	21	7
Всего	250	249	250	249	250	249
Выражен- ность при- знака	Каждый чело-век должен жить и трудится на бла-го своей Родины		Мне бы не хоте-лось, чтобы кто-нибудь вмешивался в мою жизнь		Достижение опре-делен-ного стату-са — важнейшая цель моей жизни	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	40	22	55	38	64	36
2	55	34	49	43	32	37
3	46	46	35	55	34	32
4	31	38	50	52	39	50
5	40	39	21	28	48	22
6	13	37	30	21	23	43
7	25	33	10	12	10	29
Всего	250	249	250	249	250	249

## Вариант 10

Выражен- ность при- знака	Категория		Категория		Категория	
	Юноши	Девушки	Юноши	Девушки	Юноши	Девушки
	Счастливы только богатые люди		Я не представляю свою жизнь без друзей		Важно быть уверенным в за-втрашнем дне	
1	54	67	42	67	52	56
2	48	43	31	46	41	50
3	31	30	43	35	25	42
4	37	41	36	61	43	31
5	20	24	25	35	37	33
6	28	32	26	11	25	29
7	22	23	37	5	17	19
Всего	240	260	240	260	240	260

Выраженность признака	Каждый человек должен жить и трудиться на благо своей Родины		Мне бы не хотелось, чтобы кто-нибудь вмешивался в мою жизнь		Достижение определенного статуса — важнейшая цель моей жизни	
1	61	24	45	40	61	35
2	53	35	52	32	43	41
3	37	45	30	36	32	52
4	43	54	39	44	43	62
5	19	32	35	36	30	33
6	13	43	17	33	21	25
7	14	27	22	39	10	12
Всего	240	260	240	260	240	260

## Задание № 7

**Тема:** системы ценностей в социальных группах.

**Цель:** отработать методы математического анализа систем ценностей.

**Задание.** По данным ранжирования ценностей совокупностью респондентов постройте ранжированный ряд ценностей по группе в целом и проинтерпретируйте результаты.

Порядок выполнения:

- 1) рассчитать условный индекс для каждой ценности (по формуле средней арифметической взвешенной);
- 2) проранжировать ценности по величине индекса;
- 3) рассчитать предельную ошибку выборки для определения значимости различий между ценностями с надежностью 0,95;
- 4) описать ранжированный ряд в виде текста;
- 5) сделать социологический вывод относительно полученных результатов.

Ниже представлены результаты ранжирования по значимости качеств личности студентами одного из вузов Урала.

Результаты ранжирования качеств по значимости  
(в абсолютных числах)

## Вариант 1

Ранг	Добро- та	Ум	Оптимизм	Целеу- стрем- ленность	Ак- тив- ность	Общита- тельность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	53	63	33	42	32	21	33	36	28	9
2	46	68	26	48	35	29	49	27	19	3
3	35	34	43	54	46	42	45	24	19	8
4	32	29	38	43	48	57	33	24	31	15
5	35	27	52	36	56	32	24	27	25	36

Ранг	Добро- та	Ум	Оптимизм	Целеу- стрем- ленность	Ак- тив- ность	Общительность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
6	22	27	41	37	39	24	36	32	47	45
7	34	45	23	27	31	33	28	48	29	52
8	27	26	28	25	28	47	34	49	40	46
9	38	16	30	20	21	26	44	47	51	57
10	28	15	36	18	14	39	24	36	61	79
Итого	350	350	350	350	350	350	350	350	350	350

### Вариант 2

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Общительность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	44	72	22	50	24	19	32	37	16	11
2	28	57	58	37	37	26	27	22	21	14
3	38	31	33	44	48	34	25	28	32	14
4	49	25	44	24	28	42	9	40	37	29
5	27	46	38	39	26	39	28	17	29	38
6	39	23	26	19	33	29	36	44	24	54
7	21	25	33	24	42	37	42	26	47	30
8	20	19	26	33	38	33	33	30	36	59
9	47	10	21	15	30	43	44	45	39	33
10	14	19	26	42	21	25	51	38	46	45
Итого	327	327	327	327	327	327	327	327	327	327

## Вариант 3

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	62	73	43	36	44	20	28	27	32	15
2	44	61	36	26	37	29	39	38	47	23
3	28	38	43	42	41	46	37	42	36	27
4	56	43	38	47	50	32	53	30	16	15
5	39	37	52	49	53	35	34	34	21	26
6	27	28	41	25	44	34	36	44	58	43
7	43	36	33	49	35	38	38	24	43	41
8	26	32	28	42	29	57	44	47	26	49
9	42	18	30	31	23	44	41	56	32	63
10	13	14	36	33	24	45	30	38	69	78
Итого	380	380	380	380	380	380	380	380	380	380

## Вариант 4

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	52	61	51	43	55	30	13	22	11	9
2	42	53	38	38	51	37	31	31	20	6
3	36	44	26	52	47	54	26	34	14	14
4	29	48	39	44	38	28	58	28	19	16
5	53	42	52	27	32	28	44	32	25	12
6	48	33	31	25	38	26	35	42	37	32
7	34	20	27	28	25	32	42	51	47	41
8	23	16	32	26	26	35	41	36	56	56

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
9	17	18	36	34	22	22	38	28	63	69
10	13	12	15	30	13	55	19	43	55	92
Итого	347	347	347	347	347	347	347	347	347	347

Вариант 5

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	72	82	64	41	59	32	50	17	5	8
2	53	64	67	38	42	41	43	34	42	6
3	67	49	58	57	62	29	23	45	30	10
4	46	52	63	28	38	52	42	32	49	28
5	24	41	46	56	26	48	38	39	72	40
6	42	46	38	50	22	39	45	48	37	63
7	25	28	32	28	37	39	57	59	38	87
8	42	31	37	22	45	44	53	61	39	56
9	36	24	13	46	53	52	39	49	54	64
10	23	13	12	64	46	54	40	46	64	68
Итого	430	430	430	430	430	430	430	430	430	430

Вариант 6

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	51	84	34	56	36	29	39	51	20	23
2	38	62	68	47	47	36	27	25	41	32



Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нерад- ост- ность
3	48	41	38	54	48	44	40	28	58	24
4	62	48	50	32	41	51	21	53	23	42
5	57	16	48	49	31	49	40	47	39	47
6	49	33	36	25	43	38	47	42	46	64
7	31	35	43	40	52	47	48	36	51	40
8	29	28	35	42	47	52	43	34	45	68
9	37	20	31	51	40	43	55	53	63	30
10	21	56	40	27	38	34	63	54	37	53
Итого	423	423	423	423	423	423	423	423	423	423

## Вариант 7

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нерад- ост- ность
1	65	71	61	53	65	40	23	32	21	19
2	52	66	48	48	61	47	31	51	30	16
3	46	54	39	62	57	64	36	44	24	24
4	39	58	49	57	48	38	68	38	29	26
5	63	52	62	37	45	38	54	42	35	22
6	58	43	41	35	48	39	45	52	47	42
7	44	30	37	38	35	42	55	61	57	51
8	33	26	42	36	36	45	51	49	66	66
9	27	28	46	44	32	32	48	38	76	79
10	23	22	25	40	23	65	39	43	65	105
Итого	450	450	450	450	450	450	450	450	450	450

Вариант 8

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Ак- тив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	57	63	33	26	34	10	18	17	22	5
2	34	56	26	16	27	19	29	28	37	13
3	18	28	38	32	31	36	27	32	26	17
4	46	33	28	42	40	22	43	20	6	5
5	29	27	42	39	48	25	24	24	11	16
6	17	18	31	15	34	29	26	34	48	33
7	33	26	23	39	25	28	33	14	33	31
8	16	22	18	32	19	47	34	42	16	39
9	32	8	20	21	13	34	31	46	27	53
10	3	4	26	23	14	35	20	28	59	73
Итого	285	285	285	285	285	285	285	285	285	285

Вариант 9

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Актив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	72	82	64	41	59	32	50	17	15	8
2	53	64	67	38	42	41	43	34	42	16
3	67	49	58	57	62	29	23	55	30	10
4	46	52	63	28	38	52	52	32	49	28
5	24	41	46	56	26	58	38	39	72	40
6	42	46	38	50	32	39	45	48	37	63
7	25	28	32	38	37	39	57	59	38	87
8	42	31	47	22	45	44	53	61	39	56

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Актив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
9	36	34	13	46	53	52	39	49	54	64
10	33	13	12	64	46	54	40	46	64	68
Итого	440	440	440	440	440	440	440	440	440	440

## Вариант 10

Ранг	Добро- та	Ум	Оп- ти- мизм	Целеу- стрем- ленность	Актив- ность	Об- щитель- ность	Сила	Храб- рость	На- деж- ность	Жиз- нера- дост- ность
1	53	63	33	42	32	21	33	36	33	14
2	46	68	26	48	35	29	49	27	24	8
3	35	34	48	59	46	42	45	24	19	8
4	32	29	43	48	48	57	33	24	31	15
5	35	27	52	36	61	37	24	27	25	36
6	22	27	41	37	44	29	36	32	47	45
7	34	45	23	27	31	33	33	53	29	52
8	27	26	28	25	28	47	39	54	40	46
9	43	21	30	20	21	26	44	47	51	57
10	33	20	36	18	14	39	24	36	61	79
Итого	360	360	360	360	360	360	360	360	360	360

## Задание № 8

**Тема:** статистическая оценка социологических гипотез.

**Цель:** отработать математические методы оценки и проверки гипотез.

**Задание.** По приведенным материалам социологического опроса проверьте гипотезу: предпочитаемые молодежью способы проведения досуга зависят от пола. Проверку осуществите с надежностью 0,95 и 0,99.

Порядок выполнения:

- 1) выдвинуть нулевую и альтернативную гипотезы;
- 2) определить тип критической области;
- 3) выбрать необходимый критерий проверки (статистику);
- 4) найти критические точки с необходимым уровнем значимости;
- 5) рассчитать эмпирическое значение критерия;
- 6) сравнить эмпирическое значение критерия с критическими точками;
- 7) сделать соответствующий вывод по поводу проверки гипотезы;
- 8) проверить коррелированность ранжированных рядов;
- 9) проверить значимость коэффициента ранговой корреляции Спирмена;
- 10) сделать вывод о коррелированности ранжированных рядов.

Предпочитаемые способы проведения досуга  
в зависимости от пола

### Вариант 1

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	52	113
Общение с родными	43	93
Общение с друзьями	67	82

Способы проведения досуга	Юноши	Девушки
	Частоты	
Общение в интернете	67	76
Посещение концертов	23	41
Занятия спортом	31	27
Слушание музыки	23	62
Посещение дискотек	40	106
Чтение книг, журналов	17	52
Пикники на природе	23	20
Турпоходы	14	34
Коллекционирование	4	15
Компьютерные игры	73	27

Количество ответивших: юноши — 118, девушки — 134.

### Вариант 2

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	64	78
Общение с родными	36	53
Общение с друзьями	57	47
Общение в интернете	29	39
Посещение концертов	45	28
Занятия спортом	71	34
Слушание музыки	63	53
Посещение дискотек	45	67
Чтение книг, журналов	23	32
Пикники на природе	36	17
Турпоходы	17	10
Коллекционирование	5	7
Компьютерные игры	58	36

Количество ответивших: юноши — 91, девушки — 103.

Вариант 3

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	64	126
Общение с родными	42	103
Общение с друзьями	37	96
Общение в интернете	58	78
Посещение концертов	29	45
Занятия спортом	42	32
Слушание музыки	33	74
Посещение дискотек	46	96
Чтение книг, журналов	29	57
Пикники на природе	53	45
Турпоходы	29	36
Коллекционирование	14	15
Компьютерные игры	79	37

Количество ответивших: юноши — 106, девушки — 158.

Вариант 4

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	85	122
Общение с родными	71	97
Общение с друзьями	69	84
Общение в интернете	98	76
Посещение концертов	57	67
Занятия спортом	65	53
Слушание музыки	71	69
Посещение дискотек	52	53
Чтение книг, журналов	49	61
Пикники на природе	63	53
Турпоходы	37	29

Способы проведения досуга	Юноши	Девушки
	Частоты	
Коллекционирование	13	9
Компьютерные игры	95	54

Количество ответивших: юноши — 105, девушки — 146.

### Вариант 5

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	87	123
Общение с родными	55	97
Общение с друзьями	48	84
Общение в интернете	91	76
Посещение концертов	52	69
Занятия спортом	66	53
Слушание музыки	72	72
Посещение дискотек	58	84
Чтение книг, журналов	43	57
Пикники на природе	66	42
Турпоходы	51	36
Коллекционирование	10	15
Компьютерные игры	96	28

Количество ответивших: юноши — 113, девушки — 143.

### Вариант 6

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	63	97
Общение с родными	54	83
Общение с друзьями	74	72
Общение в интернете	54	64

Способы проведения досуга	Юноши	Девушки
	Частоты	
Посещение концертов	32	31
Занятия спортом	21	47
Слушание музыки	26	72
Посещение дискотек	30	86
Чтение книг, журналов	26	52
Пикники на природе	18	20
Турпоходы	11	34
Коллекционирование	9	15
Компьютерные игры	73	57

Количество ответивших: юноши — 120, девушки — 130.

#### Вариант 7

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	87	79
Общение с родными	56	63
Общение с друзьями	47	56
Общение в интернете	59	48
Посещение концертов	34	28
Занятия спортом	67	34
Слушание музыки	53	63
Посещение дискотек	65	56
Чтение книг, журналов	13	24
Пикники на природе	46	37
Турпоходы	37	14
Коллекционирование	9	6
Компьютерные игры	59	36

Количество ответивших: юноши — 141, девушки — 123.



## Вариант 8

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	84	96
Общение с родными	52	83
Общение с друзьями	47	76
Общение в интернете	68	78
Посещение концертов	39	45
Занятия спортом	48	23
Слушание музыки	53	64
Посещение дискотек	56	86
Чтение книг, журналов	39	27
Пикники на природе	53	35
Турпоходы	49	35
Коллекционирование	12	10
Компьютерные игры	79	47

Количество ответивших: юноши — 96, девушки — 108.

## Вариант 9

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	95	112
Общение с родными	81	87
Общение с друзьями	79	74
Общение в интернете	108	66
Посещение концертов	67	57
Занятия спортом	75	43
Слушание музыки	84	59
Посещение дискотек	62	43
Чтение книг, журналов	59	51
Пикники на природе	73	43
Турпоходы	47	39

Коллекционирование	23	6
Компьютерные игры	105	64

Количество ответивших: юноши — 165, девушки — 138.

### Вариант 10

Способы проведения досуга	Юноши	Девушки
	Частоты	
Просмотр телевизора	94	126
Общение с родными	60	107
Общение с друзьями	53	94
Общение в интернете	96	86
Посещение концертов	57	79
Занятия спортом	71	63
Слушание музыки	77	82
Посещение дискотек	63	94
Чтение книг, журналов	48	67
Пикники на природе	71	52
Турпоходы	56	46
Коллекционирование	15	21
Компьютерные игры	101	48

Количество ответивших: юноши — 133, девушки — 153.

## ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ

---

1. Измерения в социологии. Виды переменных. Шкалы измерения.
2. Выборочный метод в социологии.
3. Проблема репрезентативности. Достоверность выводов социологического исследования.
4. Группировка данных. Вариационные ряды.
5. Графические методы отображения сгруппированных данных.
6. Средние величины и их смысл.
7. Надежность, точность, доверительный интервал.
8. Предельная ошибка выборки и методы ее уменьшения.
9. Статистическая проверка социологических гипотез. Алгоритм действий.
10. Проверка коррелированности ранжированных рядов данных.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

---

### *Основной*

1. Девятко, И. Ф. Методы социологического исследования / И. Ф. Девятко. — М. : Университет, 2002. — 208 с.
2. Татарова, Г. Г. Методология анализа данных в социологии / Г. Г. Татарова. — М. : Стратегия, 1998. — 222 с.
3. Толстова, Ю. Н. Анализ социологических данных / Ю. Н. Толстова. — М. : Научный мир, 2000. — 352 с.
4. Толстова, Ю. Н. Социология и математика / Ю. Н. Толстова. — М. : Научный мир, 2003. — 323 с.
5. Хили, Д. Статистика: социологические и маркетинговые исследования / Д. Хили. — СПб. : Питер; Киев: DiaSoft, 2005. — 637 с.

### *Дополнительный*

6. Девятко, И. Ф. Диагностическая процедура в социологии: Очерк истории и теории / И. Ф. Девятко. — М. : Наука, 1993. — 168 с.
7. Елисеева, И. И. Группировка, корреляция, распознавание образов : стат. методы классификации и измерения связей / И. И. Елисеева. — М. : Статистика, 1977. — 143 с.
8. Елисеева, И. И. Логика прикладного статистического анализа / И. И. Елисеева. — М. : Финансы и статистика, 1982. — 192 с.
9. Елисеева, И. И. Статистические методы измерения связей / И. И. Елисеева. — Л. : Изд-во ЛГУ, 1982. — 134 с.

10. Андреенков, В. Г. Интерпретация и анализ данных в социологических исследованиях / В. Г. Андреенков, Ю. Н. Толстова, И. И. Елисеева [и др.]. — М. : Наука, 1987. — 254 с.

11. Клигер, С. А. Шкалирование при сборе и анализе социологической информации / С. А. Клигер, М. С. Косолапов, Ю. Н. Толстова. — М. : Наука, 1978. — 112 с.

12. Андреенков, В. Г. Математические методы анализа и интерпретация социологических данных / В. Г. Андреенков, К. Д. Аргунова, В. И. Паниотто [и др.]. — М. : Наука, 1989. — 175 с.

13. Паниотто, В. И. Качество социологической информации: Методы оценки и процедуры обеспечения / В. И. Паниотто. — Киев: Наукова думка, 1986. — 206 с.

14. Паниотто, В. И. Количественные методы в социологических исследованиях / В. И. Паниотто, В. С. Максименко. — Киев : Наукова думка, 1982. — 272 с.

15. Рабочая книга социолога / под ред. Г. В. Осипова. — М. : Наука, 1983. — 477 с.

16. Андреенков, В. Г. Типология и классификация в социологических исследованиях / В. Г. Андреенков, Ю. Н. Толстова, А. А. Мирзоев [и др.]. — М. : Наука, 1982. — 296 с.

17. Толстова, Ю. Н. Логика математического анализа социологических данных / Ю. Н. Толстова. — М. : Наука, 1991. — 110 с.

18. Крымский, С. Б. Экспертные оценки в социологических исследованиях / С. Б. Крымский, Б. Б. Жилин, В. И. Паниотто [и др.]. — Киев : Наукова думка, 1990. — 318 с.

19. Ядов, В. А. Стратегия социологического исследования: Описание, объяснение, понимание социальной реальности / В. А. Ядов. — М. : Академкнига: Добросвет, 2003. — 595 с.

20. Колчинская, В. Ю. Анализ данных в социологии / В. Ю. Колчинская. — Челябинск : Изд-во ЮУрГУ, 2006. — 85 с.

## ПРИЛОЖЕНИЕ

Таблица П.1

Критические точки  $t$ -распределения Стьюдента

Число степе- ней сво- боды $k$	Уровень значимости $\alpha$ (двусторонняя критическая область)							
	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
1	3,08	6,31	12,71	31,82	63,66	127,32	318,30	636,61
2	1,89	2,92	4,30	6,96	9,92	14,09	22,33	31,60
3	1,64	2,35	3,18	4,54	5,84	7,45	10,21	12,92
4	1,53	2,13	2,78	3,75	4,60	5,60	7,17	8,61
5	1,48	2,02	2,57	3,36	4,03	4,77	5,89	6,87
6	1,44	1,94	2,45	3,14	3,71	4,32	5,21	5,96
7	1,41	1,89	2,36	3,00	3,50	4,03	4,79	5,41
8	1,40	1,86	2,31	2,90	3,36	3,83	4,50	5,04
9	1,38	1,83	2,26	2,82	3,25	3,69	4,30	4,78
10	1,37	1,81	2,23	2,76	3,17	3,58	4,14	4,59
11	1,36	1,80	2,20	2,72	3,11	3,50	4,02	4,44
12	1,36	1,78	2,18	2,68	3,05	3,43	3,93	4,32
13	1,35	1,77	2,16	2,65	3,01	3,37	3,85	4,22
14	1,34	1,76	2,14	2,62	2,98	3,33	3,79	4,14
15	1,34	1,75	2,13	2,60	2,95	3,29	3,73	4,07
16	1,34	1,75	2,12	2,58	2,92	3,25	3,69	4,02
17	1,33	1,74	2,11	2,57	2,90	3,22	3,65	3,97
18	1,33	1,73	2,10	2,55	2,88	3,20	3,61	3,92
19	1,33	1,73	2,09	2,54	2,86	3,17	3,58	3,88
20	1,33	1,72	2,09	2,53	2,85	3,15	3,55	3,85
21	1,32	1,72	2,08	2,52	2,83	3,14	3,53	3,82
22	1,32	1,72	2,07	2,51	2,82	3,12	3,51	3,79
23	1,32	1,71	2,07	2,50	2,81	3,10	3,48	3,77
24	1,32	1,71	2,06	2,49	2,80	3,09	3,47	3,75
25	1,32	1,71	2,06	2,49	2,79	3,08	3,45	3,73

Число степеней свободы $k$	Уровень значимости $\alpha$ (двусторонняя критическая область)							
	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
26	1,32	1,71	2,06	2,48	2,78	3,07	3,44	3,71
27	1,31	1,70	2,05	2,47	2,77	3,06	3,42	3,69
28	1,31	1,70	2,05	2,47	2,76	3,05	3,41	3,67
29	1,31	1,70	2,05	2,46	2,76	3,04	3,40	3,66
30	1,31	1,70	2,04	2,46	2,75	3,03	3,39	3,65
40	1,30	1,68	2,02	2,42	2,70	2,97	3,31	3,55
60	1,30	1,67	2,00	2,39	2,66	2,91	3,23	3,46
120	1,29	1,66	1,98	2,36	2,62	2,85	3,16	3,37
$\infty$	1,282	1,645	1,960	2,326	2,576	2,81	3,09	3,29
Число степеней свободы $k$	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
	Уровень значимости $\alpha$ (односторонняя критическая область)							

Таблица П.2

**Критические значения коэффициента  
ранговой корреляции Спирмена**

Количе- ство пар рангов	Вероятность ошибки			Количе- ство пар рангов	Вероятность ошибки		
	0,05	0,01	0,001		0,05	0,01	0,001
6	0,829	1,000	—	25	0,398	0,510	0,618
7	0,745	0,893	1,000	30	0,362	0,466	0,570
8	0,691	0,857	0,952	35	0,333	0,429	0,534
9	0,683	0,817	0,917	40	0,311	0,402	0,501
10	0,636	0,782	0,891	45	0,294	0,380	0,475
11	0,618	0,754	0,867	50	0,279	0,361	0,450
12	0,580	0,727	0,823	60	0,254	0,330	0,415
13	0,555	0,698	0,801	70	0,235	0,306	0,385

Количество пар рангов	Вероятность ошибки			Количество пар рангов	Вероятность ошибки		
	0,05	0,01	0,001		0,05	0,01	0,001
14	0,534	0,675	0,793	80	0,220	0,286	0,361
15	0,518	0,654	0,760	90	0,207	0,270	0,341
16	0,500	0,632	0,741	100	0,196	0,257	0,324
17	0,485	0,615	0,724	150	0,160	0,209	0,265
18	0,472	0,598	0,709	200	0,139	0,182	0,231
19	0,458	0,583	0,694	500	0,087	0,115	0,148
20	0,445	0,568	0,679	1000	0,062	0,081	0,104

Таблица П.3

Коэффициенты доверия для доверительных интервалов  
нормального распределения

$$(\text{корни уравнения } \Phi(t) = \frac{\gamma}{2}, \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz)$$

Надежность, $\gamma$	0,683	0,799	0,899	0,950	0,954	0,990	0,997	0,999
Коэффициент доверия, $t$	1,00	1,28	1,64	1,96	2,00	2,58	3,00	3,29



## Оглавление

---

ОСНОВНЫЕ ПОНЯТИЯ СТАТИСТИКИ .....	3
ФОРМУЛЫ СТАТИСТИКИ .....	8
ТИПОВЫЕ ЗАДАЧИ СОЦИОЛОГИЧЕСКОГО АНАЛИЗА .....	29
ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ.....	87
ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ .....	137
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	138
ПРИЛОЖЕНИЕ .....	140

*Учебное издание*

**Плескунов Михаил Александрович**

## **МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА СОЦИОЛОГИЧЕСКИХ ДАННЫХ**

Редактор М. А. Терновая  
Верстка Е. В. Ровнушкиной

Подписано в печать 18.01.2017. Формат 60×84 1/16.  
Бумага писчая. Цифровая печать. Усл. печ. л. 5,3.  
Уч.-изд. л. 8,4. Тираж 50 экз. Заказ 9.

Издательство Уральского университета  
Редакционно-издательский отдел ИПЦ УрФУ  
620049, Екатеринбург, ул. С. Ковалевской, 5  
Тел.: 8 (343) 375-48-25, 375-46-85, 374-19-41  
E-mail: rio@urfu.ru

Отпечатано в Издательско-полиграфическом центре УрФУ  
620075, Екатеринбург, ул. Тургенева, 4  
Тел.: 8 (343) 350-56-64, 350-90-13  
Факс: 8 (343) 358-93-06  
E-mail: press-urfu@mail.ru



